

UNIVERSIDAD NACIONAL DE PIURA

Facultad De Ciencias

**Escuela Profesional De Ingeniería Electrónica y
Telecomunicaciones**



TESIS

**“SISTEMA DE LECTURA AUTOMÁTICA DEL NUMERO DE DNI
UTILIZANDO VISION ARTIFICIAL”**

PRESENTADA POR:

Br. Esneider Alexander Neira Mija

ASESOR: Mg. JUAN M. JACINTO SANDOVAL

**PARA OPTAR EL TITULO PROFESIONAL DE
INGENIERO ELECTRÓNICO Y TELECOMUNICACIONES**

Línea de investigación: Informática, Electrónica y Telecomunicaciones

Sub – Línea: Sistemas Digitales

Piura, Perú

2017

UNIVERSIDAD NACIONAL DE PIURA

Facultad De Ciencias

Escuela Profesional De Ingeniería Electrónica y Telecomunicaciones



“SISTEMA DE LECTURA AUTOMÁTICA DEL NUMERO DE DNI UTILIZANDO VISION ARTIFICIAL”

TESIS

Para optar el título profesional de:

INGENIERO ELECTRONICO Y TELECOMUNICACIONES

Mg. Juan Manuel Jacinto Sandoval
ASESOR

Br. Esneider Alexander Neira Mija
AUTOR

Piura – Perú
Año 2017

DECLARACION JURADA DE AUTENTICIDAD DE LA TESIS

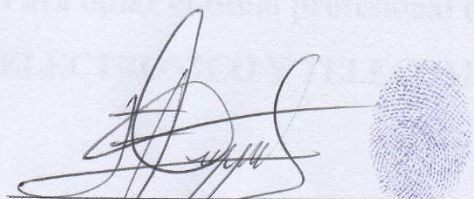
Yo: Esneider Alexander Neira Mija con DNI N° 43513761, Bachiller de la escuela Profesional de Ingeniería Electrónica y Telecomunicaciones, de la Facultad de Ciencias domiciliado en A.H San Sebastián Mz B7 lote 11 distrito 26 de octubre, Piura.

Celular: 957950553

Email: esneider.mija@gmail.com

DECLARO BAJO JURAMENTO: que la tesis que presento es autentica e inédita, no siendo copia parcial ni total de una tesis desarrollada, y/o realizada en el Perú o en el Extranjero, en caso contrario de resultar falsa la información que proporcione, me sujeto a los alcances de lo establecido en el Art. N° 411, del código penal concordante con el Art. 32° de la ley N° 27444, y la ley del Procedimiento Administrativo General y las Normas Legales de Protección a los Derechos de Autor.

Piura, mayo del 2017



Esneider Alexander Neira Mija

Artículo 411.- el que, en un procedimiento administrativo, hace una falsa declaración en relación con hechos o circunstancias que le corresponde probar, violando la presunción de veracidad establecida por ley, será reprimido con pena privativa de libertad no menor de uno ni mayor de cuatro años.

Art. 4. Inciso 4.12 del Reglamento del Registro Nacional de trabajos de Investigación para optar grados académicos y títulos profesionales – RENATI Resolución de Consejo Directivo N° 033-2016-SUNEDU/CD.

UNIVERSIDAD NACIONAL DE PIURA

Facultad De Ciencias

**Escuela Profesional De Ingeniería Electrónica y
Telecomunicaciones**



“SISTEMA DE LECTURA AUTOMÁTICA DEL NUMERO DE DNI UTILIZANDO VISION ARTIFICIAL”

TESIS

Para optar el título profesional de:

INGENIERO ELECTRONICO Y TELECOMUNICACIONES

Dr. Carlos Enrique Arellano Ramírez
PRESIDENTE

Ing. Franklin Barra Zapata
SECRETARIO

Ing. Miguel Ángel Panduro Alvarado
VOCAL

PIURA – PERU

2017



UNIVERSIDAD NACIONAL DE PIURA

FACULTAD DE CIENCIAS



ACTA DE SUSTENTACIÓN 052-2017-FC-UNP

FACULTAD DE CIENCIAS

Los Miembros del Jurado Calificador que suscriben, reunidos para evaluar la Tesis denominada **"SISTEMA DE LECTURA AUTOMÁTICA DEL NÚMERO DE DNI UTILIZANDO VISIÓN ARTIFICIAL"** presentado por el señor Bachiller **NEIRA MIJA ESNEIDER ALEXANDER**, con el asesoramiento del **MSc. Juan Manuel Jacinto Sandoval**; oídas las observaciones y respuestas a las preguntas formuladas, y de conformidad al Reglamento de Tesis para obtener el Título Profesional en la Facultad de Ciencias, lo declaran:

APROBADO (X)

DESAPROBADO ()

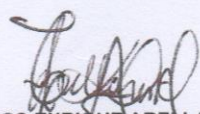
Con la mención de:

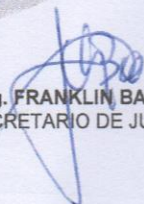
..... MUY BUENO

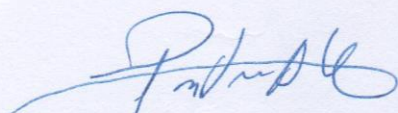
(X) En consecuencia, queda en condición de ser ratificado por el Consejo de Facultad de Ciencias de la Universidad Nacional de Piura, y recibir el **TÍTULO PROFESIONAL DE INGENIERO ELECTRÓNICO Y TELECOMUNICACIONES**.

(X) En consecuencia, queda en condición de ser ratificado por el Consejo Universitario de la Universidad Nacional de Piura, y recibir el **TÍTULO PROFESIONAL DE INGENIERO ELECTRÓNICO Y TELECOMUNICACIONES**; después que el sustentante incorpore la sugerencia del Jurado Calificador.

Piura, 21 de julio de 2017.


Dr. CARLOS ENRIQUE ARELLANO RAMÍREZ
PRESIDENTE DE JURADO DE TESIS


Ing. FRANKLIN BARRA ZAPATA
SECRETARIO DE JURADO DE TESIS


Ing. MIGUEL ÁNGEL PANDURO ALVARADO
VOCAL DE JURADO DE TESIS



DEDICATORIA

Mi tesis se la dedico en primer lugar a DIOS por ser parte de magnificencia a mis queridos padres por todo lo bueno inculcado en mi persona por lo grandiosos que son, por la inmensa paciencia que han tenido y tienen aún.

A mis familiares hermanos tíos y amigos por el apoyo incondicional para ser mejor persona y profesional.

A mis maestros por todos sus conocimientos y experiencia que nos brindaron para tener las habilidades requeridas y desempeñarnos como excelentes profesionales.

AGRADECIMIENTO

Este trabajo de tesis primer lugar me gustaría agradecerle a ti Dios por bendecirme para llegar hasta donde he llegado, porque hiciste realidad este objetivo.

A mi asesor de tesis, Ing. Juan Manuel Jacinto Sandoval por su esfuerzo y dedicación quien, con sus conocimientos, su experiencia, su paciencia y su motivación ha logrado en mí que pueda terminar mis estudios con éxito.

También me gustaría agradecer a mis profesores durante toda mi carrera profesional porque todos han aportado con un granito de arena a mi formación.

Son muchas personas que han formado parte de mi vida profesional a las que me encantaría agradecerles su amistad, consejos, apoyo, ánimo y compañía en los momentos más difíciles de mi vida. Algunas están aquí conmigo y otras en mi recuerdo y en mi corazón, sin importar en donde estén quiero darles las gracias por formar parte de mí, por todo lo que me han brindado y por todas sus bendiciones.

A todos ustedes, mi mayor reconocimiento y gratitud.

RESUMEN

En la presente Tesis, se diseña e implementa un sistema de lectura automática del número del Documento Nacional de Identidad, DNI, utilizando visión artificial. Para eso se hace uso de una cámara webcam y una computadora personal, como hardware y lenguaje de programación C++, con librerías de visión artificial, como software.

Se toma como entrada la captura de una imagen con fondo de color blanco, donde está el DNI, teniendo en cuenta la iluminación de la escena. Esta imagen se captura en color, para luego aplicarle algoritmos de mejoramiento, segmentación y binarización. Una vez que se segmenta el DNI, que es de color celeste, se procede a localizar la parte del número, que está por debajo de la foto de la persona. Ahora, el último paso es la etapa del reconocimiento de los 8 dígitos.

Visión artificial, dni, imágenes, lectura automática.

ABSTRACT

In this thesis, a system of automatic reading of the National Identity Document number, DNI, using artificial vision is designed and implemented. For that it makes use of a webcam and a personal computer, like the hardware and the programming language C ++, with libraries of artificial vision, like software.

It is taken as the capture of an image with white background, where is the DNI, taking into account the lighting of the scene. This image is captured in color, then applied to improvement, segmentation and binarization algorithms. Once it is separated from the DNI, which is light blue, proceed to locate the part of the number, which is below the photo of the person. Now, the last step in the stage of recognizing the 8 digits.

Artificial vision, dni, images, automatic reading

INDICE GENERA

INTRODUCCIÓN.....	1
I. ASPECTOS DE LA PROBLEMÁTICA.....	2
1.1. DESCRIPCIÓN DE LA REALIDAD PROBLEMÁTICA.	2
1.2. JUSTIFICACIÓN E IMPORTANCIA DE LA INVESTIGACIÓN.....	2
1.3. OBJETIVOS.....	3
1.4. DELIMITACIÓN DE LA INVESTIGACIÓN	3
II. MARCO TEORICO.....	4
2.1. ANTECEDENTES DE LA INVESTIGACIÓN	4
2.2. BASES TEÓRICAS	4
2.2.1. <i>DOCUMENTO NACIONAL DE IDENTIDAD</i>	4
2.2.2. <i>HISTORIA</i>	5
2.3. CARACTERÍSTICAS DEL DNI ELECTRÓNICO	8
2.3.1. <i>CARACTERISTICAS DEL MICROCHIP</i>	8
2.3.2. <i>CARACTERISTICAS DE LA TARJETA ELECTRONICA</i>	8
2.3.3. <i>INFORMACION MOSTRADA EN EL DNI ELECTRONICO</i>	9
2.3.4. <i>INTELIGENCIA ARTIFICIAL</i>	10
2.3.5. <i>VISION ARTIFICIAL</i>	10
2.3.6. <i>PROCESAMIENTO DE IMAGENES</i>	11
2.3.7. <i>LA IMAGEN DIGITAL</i>	11
2.3.8. <i>DISPOSITIVOS DE CAPTURA DE IMÁGENES</i>	12
2.3.9. <i>IMÁGENES BLANCO/NEGRO Y COLOR</i>	15
2.3.10. <i>RESOLUCIÓN ESPACIAL Y EN AMPLITUD</i>	16
2.3.11. <i>REPRESENTACIÓN DE IMÁGENES DIGITALES</i>	17
2.4. PROCESAMIENTO Y ANALISIS DE IMÁGENES DIGITALES.....	18
2.5. PROCESAMIENTO BASICO DE IMÁGENES.....	18
2.5.1. <i>OPERACIONES INDIVIDUALES</i>	19
2.5.2. <i>OPERADOR IDENTIDAD</i>	20
2.5.3. <i>OPERADOR INVERSO O NEGATIVO</i>	21
2.5.4. <i>OPERADOR UMBRAL</i>	22
2.5.5. <i>TRANSFORMACIÓN DE VECINDAD</i>	23

2.6.	OPERACIONES DE FILTRADO	25
2.6.1.	<i>FILTROS PASA BAJO</i>	25
2.6.2.	<i>FILTROS PASA ALTO</i>	25
2.7.	HISTOGRAMA	26
2.8.	ESPACIOS DE COLORES Y COLORIMETRÍA	27
2.8.1.	<i>ESPACIO DE COLOR RGB</i>	28
2.8.2.	<i>ESPACIO DE COLOR HSV</i>	29
2.9.	TESSERACT OCR.....	31
2.9.1.	<i>HISTORIA</i>	31
2.9.2.	<i>CARACTERÍSTICAS</i>	31
2.10.	GLOSARIO DE TÉRMINOS BÁSICOS	31
2.11.	HIPÓTESIS.....	34
III.	MARCO METODOLÓGICO	35
3.1.	MÉTODOS Y PROCEDIMIENTOS.....	35
3.1.1.	<i>CAPTURA</i>	35
3.1.2.	<i>RGB A HSV</i>	36
3.1.3.	<i>DEFINIR EL RANGO DEL COLOR CELESTE</i>	36
3.1.4.	<i>BINARIZAR DNI</i>	36
3.1.5.	<i>ETIQUETAR OBJETOS</i>	36
3.1.6.	<i>SEGMENTAR EN BASE AL AREA</i>	36
3.1.7.	<i>DIVIDIR EN 3 REGIONES HORIZONTALES EL DNI</i>	37
3.1.8.	<i>BINARIZACIÓN DE 3° REGIÓN A PARTIR DE LA IMAGEN ORIGINAL</i> 37	
3.1.9.	<i>ETIQUETAR OBJETOS EN LA 3ERA REGION DE INTERÉS</i>	37
3.1.10.	<i>BUSCAR EL CARÁCTER</i>	37
3.1.11.	<i>LECTURA DEL DNI</i>	37
IV.	RESULTADOS Y DISCUSIÓN.....	38
4.1.	RESULTADOS	38
4.2.	DISCUSIÓN.....	46
	CONCLUSIONES	48
	RECOMENDACIONES	49

REFERENCIAS BIBLIOGRÁFICAS	50
ANEXOS	51

ÍNDICE DE FIGURAS

FIGURA 1:DOCUMENTO NACIONAL DE IDENTIDAD	8
FIGURA 2: DIGITALIZACIÓN DE UNA SEÑAL ANALÓGICA	13
FIGURA 3: CAPTURA DE UNA IMAGEN 3D POR UN DISPOSITIVO CCD	14
FIGURA 4: FIGURA DEL ÁRBOL CAPTURADA POR UNA CÁMARA CON 4X4 SENSORES DE INTENSIDAD	15
FIGURA 5: CUATRO REPRESENTACIONES DE LA MISMA IMAGEN CON VARIACIÓN EN EL NÚMERO DE PÍXELES UTILIZADOS	16
FIGURA 6: SEIS REPRESENTACIONES DE LA MISMA IMAGEN CON VARIACIÓN EN EL NÚMERO DE NIVELES DE GRIS UTILIZADOS.....	17
FIGURA 7: CONVENCIÓN DE EJES UTILIZADA PARA LA REPRESENTACIÓN DE IMÁGENES DIGITALES.....	18
FIGURA 8: FUNCIONES DE PUNTO Y VECINDAD.....	19
FIGURA 9: OPERACIÓN INDIVIDUAL	20
FIGURA 10: REPRESENTACIÓN DEL OPERADOR IDENTIDAD.....	21
FIGURA 11: REPRESENTACIÓN DEL OPERADOR INVERSO	22
FIGURA 12: REPRESENTACIÓN DEL OPERADOR UMBRAL.....	22
FIGURA 13: TRES NÚCLEOS REPRESENTATIVOS DE FILTROS PASO BAJO...	25
FIGURA 14: TRES NÚCLEOS REPRESENTATIVOS DE FILTROS PASO ALTO..	26
FIGURA 15: HISTOGRAMA	26
FIGURA 16: MODELO ADITIVO RGB	29
FIGURA 17: PESO DE LAS 3 COMPONENTES EN EL MODELO RGB.....	29
FIGURA 18: MODELO HSV	30
FIGURA 19: TINTE	30
FIGURA 20: SATURACIÓN	30
FIGURA 21: VALOR	30
FIGURA 22: IMAGEN DE ENTRADA CON DNI	38
FIGURA 23: IMAGEN SEGMENTADA POR COLOR Y AREA.....	39
FIGURA 24: IMAGEN BINARIZADA	39
FIGURA 25: REGION DONDE SE ENCUENTRA EL NÚMERO	40
FIGURA 26: IMAGEN CON EL RESULTADO DE LA LECTURA.....	40
FIGURA 27: IMAGEN DE ENTRADA CON DNI	41

FIGURA 28: IMAGEN SEGMENTADA POR COLOR Y AREA	41
FIGURA 29: IMAGEN BINARIZADA	42
FIGURA 30: REGION DONDE SE ENCUENTRA EL NÚMERO	42
FIGURA 31: IMAGEN CON EL RESULTADO DE LA LECTURA.....	42
FIGURA 32: IMAGEN DE ENTRADA CON DNI	43
FIGURA 33: IMAGEN BINARIZADA	43
FIGURA 34: REGION DONDE SE ENCUENTRA EL NÚMERO	44
FIGURA 35: IMAGEN CON EL RESULTADO DE LA LECTURA.....	44
FIGURA 36: IMAGEN DE ENTRADA CON DNI	45
FIGURA 37: IMAGEN BINARIZADA	45
FIGURA 38: REGION DONDE SE ENCUENTRA EL NÚMERO	46
FIGURA 39: IMAGEN CON EL RESULTADO DE LA LECTURA.....	46

ÍNDICE DE ANEXOS

ANEXO 1: CÓDIGO EN C++ DEL PROGRAMA.....	51
ANEXO 2: RESULTADOS DE LECTURA DE IMÁGENES IMPRESAS DE DNI DE INTERNET.....	60

INTRODUCCIÓN

Es preciso reconocer que hoy por hoy la visión por computador a veces no es la mejor solución a un problema. Existen muchas ocasiones en las que el problema es tan complejo que la solución humana es lo mejor. Por ejemplo, imaginemos una conducción de un vehículo en una carretera con tráfico intenso. Pero a veces, las soluciones humanas tienden a ser inexactas o subjetivas y en ocasiones lentas y presentan una ausencia de rigor, así como una pobre percepción. No obstante, la solución humana es menos estructurada que la solución artificial y muchos problemas de visión por computador requieren un nivel de inteligencia mucho mayor que el que la máquina pueda ofrecer. El sistema de visión humana puede describir automáticamente una textura en detalle, un borde, un color, una representación bidimensional de una tridimensional, ya que puede diferenciar entre imágenes de diferentes personas, firmas, colores, etc., puede vigilar ciertas zonas, diagnosticar enfermedades a partir de radiografías, etc. Sin embargo, aunque algunas de estas tareas pueden llevarse a cabo mediante visión artificial, el software o el hardware necesario no consigue los resultados que serían deseables. Aun a pesar de las limitaciones expuestas, cada día es mayor el número de aplicaciones de la visión artificial. Por ello sólo es posible dar una pequeña pincelada sobre las múltiples aplicaciones en las que la visión artificial se ha aplicado hasta el momento.

I. ASPECTOS DE LA PROBLEMÁTICA

1.1. DESCRIPCIÓN DE LA REALIDAD PROBLEMÁTICA.

El número del documento del DNI, como ya sabemos identifica a las personas peruanas. Existen dispositivos tecnológicos, que leen un código que tiene el DNI, pero estos funcionan con una computadora al costado. Este código es el llamado código de barras, que son un conjunto de barras blancas y negras, que tienen un significado numérico, que representa el número de identificación de una persona.

Lo que se pretende con este proyecto de tesis, es resolver la lectura del número de identificación, mediante reconocimiento óptico de caracteres. Éste sería el primer paso, para poder posteriormente en un trabajo futuro, reconocer los apellidos y nombres de la persona y que sea implementada en un celular, para que agilice la lectura, por ejemplo, para ayudar a las autoridades policiales, cuando se necesita saber si una persona tiene requisitoria y no estar llamando por celular a la central, pidiendo información del individuo.

¿Será posible implementar un sistema de visión artificial para reconocer el número del documento nacional de identidad?

1.2. JUSTIFICACIÓN E IMPORTANCIA DE LA INVESTIGACIÓN

Para la lectura automática del número del DNI, existen dispositivos en el mercado, que hacen esta tarea, pero para ello necesitan una computadora personal al lado. La policía nacional del Perú, constantemente, como parte de su trabajo hacen requisitorias, donde, por ejemplo, detienen a los vehículos y comienzan a pedir el DNI a todos los ocupantes. En los vehículos grandes, como los ómnibus, se emplean demasiado tiempo, ya que comienzan a pedir los DNI de todas las personas, para posteriormente, llamar por teléfono y dictar uno por uno, y preguntar si alguno es requisitoriado. Del otro lado hay una persona que verifica si lo es. Esto se puede solucionar utilizando el mismo celular y su cámara a través de visión artificial. La lectura y el procesamiento de verificación en tiempo real. Para comenzar, se hará este proyecto leyendo el número del DNI utilizando visión artificial en una computadora personal.

1.3. OBJETIVOS

OBJETIVO GENERAL

Utilizando visión artificial, se reconocerá automáticamente el número del documento nacional de identificación de una persona.

OBJETIVOS ESPECÍFICOS

- Implementar técnicas de procesamiento para mejorar la imagen del DNI.
- Implementar algoritmos de reconocimiento de caracteres.
- Implementar el sistema utilizando una cámara webcam.

1.4. DELIMITACIÓN DE LA INVESTIGACIÓN

El proyecto será desarrollado, en condiciones ambientales de iluminación y teniendo en cuenta la posición del Documento Nacional de Identidad, es decir no se considerará la rotación del DNI, además se hará teniendo en cuenta una sola distancia de la cámara con respecto del DNI.

El DNI que se tomará en cuenta para el propósito de esta tesis son los de color celeste. El proyecto se desarrollará en una computadora personal de características básicas.

II. MARCO TEORICO

2.1. ANTECEDENTES DE LA INVESTIGACIÓN

En la tesis “Reconocimiento de caracteres mediante imágenes en contadores de gas en entornos reales” cuyos autores, Ignacio Hidalgo Bejarano, Roberto José Sánchez y García de Blas, Madrid, junio de 2015, donde el objetivo fundamental del proyecto era poder resolver el problema planteado por la empresa Madrileña Red de Gas, que era la lectura del consumo actual del contador, así como su número de serie. La tasa de acierto pretendida era sobre imágenes encuadradas conteniendo el contador, en entornos reales. Esto supone un paso muy importante para el usuario final, ya que éste no tiene que estar pendiente de encuadrar sobre la imagen la parte relativa a la parte frontal del contador.

2.2. BASES TEÓRICAS

2.2.1. DOCUMENTO NACIONAL DE IDENTIDAD¹

El Documento Nacional de Identidad (DNI) del Perú constituye la única cédula de identidad personal reconocida por el Estado para todos los actos civiles, comerciales, administrativos, judiciales y, en general, para todos aquellos casos en que, por mandato legal, deba ser presentado. Es un documento público, personal e intransferible y constituye también el único título de derecho al sufragio de la persona a cuyo favor ha sido otorgado. Su emisión se encuentra a cargo del Registro Nacional de Identificación y Estado Civil - RENIEC.

A partir del 15 de julio de 2013 el RENIEC viene expidiendo el DNI electrónico (DNI-e) ¹ que reemplazará gradualmente al DNI actual. El DNI electrónico está fabricado en policarbonato y tiene el formato de una tarjeta de crédito, siguiendo la norma ISO 7816. Posee un chip basado en las tecnologías de firma electrónica, tarjeta inteligente y biométrica, e incorpora inicialmente cuatro aplicativos de software: El primero de identidad eMRTD ICAO, el segundo de firma digital PKI, el tercero de autenticación biométrica por huella dactilar Match-on-Card y un cuarto de tipo genérico que incluya el almacenamiento de datos y dispositivos

¹ [https://es.wikipedia.org/wiki/Documento_Nacional_de_Identidad_\(Per%C3%BA\)](https://es.wikipedia.org/wiki/Documento_Nacional_de_Identidad_(Per%C3%BA))

contadores. En junio de 2015, el DNI electrónico fue reconocido como el mejor documento de identidad de América Latina, durante la «Conferencia Latinoamericana de Imprenta de Alta Seguridad» efectuada en Lima, la cual estuvo organizada por la firma británica Reconnaissance International, dedicada a holografía, moneda, autenticación y seguridad documentaria.

La validez del DNI es de ocho años, plazo a cuyo término los ciudadanos tienen la obligación de realizar el respectivo trámite de renovación (de ser necesario modificar algunos datos, estos tienen que realizarse sí o sí). Esto con la finalidad de mantener actualizados los datos en el registro civil. Sin embargo, cuando una persona renueva su DNI a los 70 años o más, este ya no caducará, debido a que la persona lo renovó en una edad en la que ya no es obligatorio el tener que ir a votar en las elecciones; sin embargo, se recomienda renovarlo en caso de modificar algunos datos como el estado civil (en caso de enviudar o divorciarse), el cambio de domicilio o de número de teléfono, etc.

2.2.2. HISTORIA

Diversos han sido los sistemas y documentos utilizados en el Perú con la finalidad de llevar un registro de su población. Al inicio estos sistemas sólo tenían como finalidad la contabilización de los nacimientos y defunciones, para posteriormente irse complejizando hasta registrar todos los actos civiles, así como dotar a la población de un documento de identidad.

2.2.2.1. Durante el Imperio Inca

Durante el Imperio Inca se procuraba registrar todos los hechos de importancia para la sociedad, por lo cual se contaban los nacimientos, las defunciones y otros eventos. Según el Inca Garcilaso de la Vega, los incas “utilizaban el entrelazado de cintas de colores y nudos” para realizar sus registros, a los cuales denominaba “quipus” y que estaban a cargo de un administrador inca, al que llamaban “Quipucamayoc”. Los registros incluían los nacimientos y muertes por cada mes, los tributos pagados al Inca cada año, especificando cada hogar que así lo hacía, así como la totalidad de personas que iban a la guerra y morían.

2.2.2.2. Durante la época colonial

Durante la época colonial la Iglesia Católica cumplía las funciones registrales a través de las parroquias, rigiéndose por el Derecho Canónico. Mediante Real

Orden del 21 de marzo de 1749 se ordenó la formación de estados mensuales de los nacimientos, matrimonios y defunciones, encargándoles el cuidado y custodia de estos libros en las mismas parroquias.

2.2.2.3. *Durante la República*

Luego de producida la independencia del Perú, el Estado se hace cargo de dichas funciones delegándose en una primera fase esas responsabilidades a las prefecturas, subprefecturas y gobernaciones.

El 21 de junio de 1852 se promulga el primer Código Civil del Perú en el cual se crean los Registros del Estado Civil, cuya gestión estaba a cargo de los gobernadores distritales. Por la Constitución de 1856, son los alcaldes provinciales y distritales los encargados de administrar el registro.

2.2.2.4. *Creación de la Libreta Electoral*

En 1931 se crea el Registro Electoral en el Perú y se entrega por primera vez la Libreta Electoral (LE) a ciudadanos hábiles para votar (hombres mayores de 21 años que supieran leer y escribir). También servía como cédula de identidad y fue la primera en el país, pero su principal propósito era para fines electorales.

En 1946 se crea la Libreta Electoral Municipal (LEM). Este documento tenía la apariencia de un pequeño libro con tapa y contratapa y cuatro hojas interiores. El material del cual estaba elaborado era papel simple. Entre las medidas de seguridad se encontraban la huella digital, la firma del registrador, firma de inscrito y el sello de Registro Electoral Municipal.

En 1963 se crea la Libreta Electoral de 7 Dígitos que era un documento de tres cuerpos. Este documento se entregó durante el primer gobierno del presidente Fernando Belaunde Terry. El material en el que se elaboró esta libreta era cartulina dúplex delgada de color blanco.

En 1978 se crea el Registro de Identidad para Analfabetos (RIA), pues antes de la aparición de este documento los analfabetos no eran tomados en cuenta para ser identificados. Este documento estaba elaborado en cartulina simple de color rosa, y constaba de dos cuerpos. Entre las medidas de seguridad mencionamos la parte anverso con la inscripción en letras en miniatura con el texto “Decreto Ley N° 22379”, sello y firma del registrador, la impresión de huella dactilar, y el sello del Registro de Identificación del Jurado Nacional de Elecciones encima de la fotografía.

En 1984 se crea la Libreta Electoral de 8 Dígitos. El material que se siguió utilizando en estos años fue la cartulina dúplex delgada de color blanco a tres cuerpos. Entre sus medidas de seguridad tenía la inscripción en color salmón con letras en miniatura el texto “Jurado Nacional de Elecciones”, número de la Libreta Electoral en la portada, firma del inscrito, sello y firma del registrador, sello del Jurado Nacional de Elecciones encima de la fotografía e impresión dactilar.

2.2.2.5. Creación del Documento Nacional de Identidad (DNI)

En 1993 se crea el Registro Nacional de Identificación y Estado Civil (RENIEC) como organismo autónomo del Estado encargado de registrar los nacimientos, matrimonios, defunciones, divorcios y otros que modifican el estado civil. Posteriormente, en 1995 se crea el Registro Único de Identificación de las Personas Naturales.

En 1996 a 1998 se crea la Libreta Electoral Mecanizada (LEM), documento de carácter transitorio pero cuya información ya era digitada por medio computarizado, no existiendo dato alguno que fuera inscrito manualmente.

En 1997 se crea el Documento Nacional de Identidad (DNI) como cédula de identidad personal en reemplazo definitivo de la Libreta Electoral. Documento Público, personal e intransferible. Su formato original fue ISO ID-02 de 10 cm de ancho por 7 cm de largo y de un solo cuerpo. En el 2005 se aprueba el formato ISO ID-01 similar al tamaño de una tarjeta de crédito 8.54 cm ancho por 5.4 cm de largo.

Con la creación del DNI también se crea el DNI de menores cuya impresión es de diferente color al de adultos. Las medidas son las mismas del DNI adulto a excepción de la carencia de huella digital y firma escaneada que corresponde al adulto que lo declara ante el RENIEC. Formato original ISO ID-02, 10 cm de ancho por 7 cm largo.

En el 2005 se aprueba el formato ISP ID-01 de tamaño 8.54 cm de ancho × 5.4 cm de largo. El DNI de menores sirve para asegurar el acceso a los servicios de salud, educación, alimentación y seguridad a los niños y adolescentes peruanos.

En 2013 se implementa el DNI Electrónico, documento que posee un chip que permite al ciudadano identificarse en Internet y “hacer transacciones con el Estado desde cualquier punto de acceso a la red, por ejemplo, desde la comodidad de su

casa”. Los jóvenes de 18 años que por primera vez tramiten su DNI serán los primeros en obtenerlo en su versión electrónica.



FIGURA 1:DOCUMENTO NACIONAL DE IDENTIDAD

2.3. CARACTERÍSTICAS DEL DNI ELECTRÓNICO

2.3.1. CARACTERISTICAS DEL MICROCHIP

- Sistema operativo Java Card, que posibilita la incorporación de futuras aplicaciones y contenidos.
- Capacidad criptográfica para gestión de claves RSA y firma digital con certificados.
- Memoria EEPROM de 144 Kb.
- Seguridad conforme a los estándares internacionales Common Criteria nivel EAL4+ ó FIPS 140-2.
- Basic Access Control (BAC), que previene el acceso no autorizado al contenido del chip.
- Active Authentication (AA), clave RSA de 1024 bits que garantiza la autenticidad del chip e impide su clonación.
- Aplicaciones: PKI, ICAO eMRTD, Match On Card (MOC).
- Software complementario: Middleware, SDK para aplicaciones cliente, Java Card SDK.

2.3.2. CARACTERISTICAS DE LA TARJETA ELECTRONICA

- Fondo con patrón Guilloché
- Impresión de textos o motivos visibles con luz ultravioleta

- Micro línea offset
- Zona de foto con micro texto ondulado
- Dispositivo Ópticamente Variable (DOVID)
- Grabado Láser
- Tinta ópticamente variable (Optical Variable Ink - OVI)
- Elemento de seguridad microscópico (JDSU charms)
- Imagen Láser Variable (CLI)

2.3.3. INFORMACION MOSTRADA EN EL DNI ELECTRONICO

De forma similar al DNI anterior, el DNI-e contiene la siguiente información:

- Número de CUI (Cédula Única de Identidad)
- Primer apellido (apellido paterno)
- Segundo apellido (apellido materno)
- Prenombres (nombres)
- Sexo
- Estado civil
- Fecha de nacimiento
- Ubigeo de nacimiento (código del departamento, provincia y distrito)
- Fecha de emisión
- Fecha de caducidad
- Grupo de votación (número de mesa de votación)
- Donación de órganos
- Fotografía
- Firma
- Huella del índice derecho
- Departamento, provincia y distrito de domicilio
- Dirección
- Número de teléfono fijo

2.3.4. INTELIGENCIA ARTIFICIAL²

A pesar de que la mayoría de los intentos para definir términos complejos y a la vez ampliamente usados suelen ser inútiles, es positivo al menos esbozar los límites aproximados en los que encuadrar el concepto de la IA para poder proporcionar una perspectiva a la explicación que sigue a continuación. Para lograrlo, se ha propuesto la siguiente definición, a pesar de no ser aceptada universalmente. La Inteligencia Artificial (IA) estudia cómo lograr que las máquinas realicen tareas que, por el momento, son realizadas mejor por los seres humanos.

Esta definición es, por supuesto, bastante efímera ya que hace referencia al estado actual de la informática.

2.3.5. VISION ARTIFICIAL

También conocida como visión por computador (del inglés computer vision) o visión técnica, es un subcampo de la inteligencia artificial y es el campo de acción más ambicioso del procesamiento digital de imágenes. Básicamente el objetivo es automatizar funciones de inspección visual, tradicionalmente utilizadas por el hombre como, por ejemplo.

Los objetivos típicos de la visión artificial incluyen:

La detección, segmentación, localización y reconocimiento de ciertos objetos en imágenes (por ejemplo, caras humanas).

La evaluación de los resultados (ej.: segmentación, registro).

Registro de diferentes imágenes de una misma escena u objeto, hacer concordar un mismo objeto en diversas imágenes.

Seguimiento de un objeto en una secuencia de imágenes.

Mapeo de una escena para generar un modelo tridimensional de la escena; tal modelo podría ser usado por un robot para navegar por la escena.

Estimación de las posturas tridimensionales de humanos.

² Elaine Rich, Kevin Knight, "Inteligencia Artificial"

Búsqueda de imágenes digitales por su contenido.

2.3.6. PROCESAMIENTO DE IMAGENES³

Imagen: Es la proyección en perspectiva en el plano bidimensional de una escena tridimensional en un determinado instante de tiempo t_0 .

Fotograma: Es una matriz bidimensional de valores de intensidad lumínica obtenidos para un tiempo t_0 constante. Podiera decirse en cierta forma que es una imagen discretizada.

Píxel (Picture Element): Es cada de una de las posiciones en que es discretizada una imagen, o lo que es lo mismo, cada una de las posiciones de un cuadro.

Imagen binaria: Son aquellas imágenes cuyos píxeles solo tienen dos valores: cero y uno.

2.3.7. LA IMAGEN DIGITALA

Son el principal ingrediente de lo que se conoce como Visión Artificial y representan mediante algún tipo de codificación, normalmente en una matriz de números de dos dimensiones, una escena del entorno.

Existen dos tipos de imágenes utilizadas frecuentemente en Visión Artificial: imágenes de intensidad e imágenes de alcance (también llamadas imágenes de profundidad o perfiles de superficie). Las imágenes de intensidad miden la cantidad de luz que incide en un dispositivo fotosensible, mientras que las imágenes de alcance estiman directamente la estructura en tres dimensiones (3D) de la escena ya que su fundamento radica en el uso de sensores de alcance ópticos y algún fenómeno físico para adquirir la imagen. Un ejemplo típico de una imagen de intensidad es una fotografía, mientras que de una imagen de alcance es, por ejemplo, la imagen que obtiene el oftalmólogo sobre el grado de rugosidad de la córnea de un paciente o las imágenes de un radar.

³ Rafael C. Gonzales, Richard E. Woods, “Tratamiento Digital de Imágenes”

⁴ Gonzalo Pajares Martinsanz, Jesús Manuel de la Cruz García, José Manuel molina pascual, Juan Cuadrado Pardo, Alejandro López Correa, “Imágenes digitales - Procesamiento practico con java”

Aunque la filosofía de los diferentes tipos de imágenes es diferente, en cualquier caso, tras su captura tendremos una matriz de valores en dos dimensiones (2D), es decir, una imagen digital.

2.3.8. *DISPOSITIVOS DE CAPTURA DE IMÁGENES*

Para la adquisición de imágenes digitales se requieren dos elementos básicos. El primero es un dispositivo físico que es sensible a una determinada banda del espectro de energía electromagnético (tal como rayos X, ultravioleta, visible, infrarrojo, etc.) y que produce una señal eléctrica de salida proporcional al nivel de energía incidente en cualquier instante de tiempo. El segundo, denominado digitalizador, es un dispositivo que cumple la función de convertir la señal eléctrica continua de salida del dispositivo físico en un conjunto discreto de localizaciones del plano de la imagen y, después, en la cuantización de dicha muestra. Esto implica, en primer lugar, determinar el valor de la imagen continua en cada una de las diferentes localizaciones discretas de la imagen (cada valor localizado de forma discreta se denomina muestra de la imagen) y, luego, asignar a cada muestra una etiqueta entera discreta, que es representativa del rango en el que varía la muestra.

Una vez capturada la señal continua y cuantificada espacialmente y en amplitud, se obtiene una imagen digital, que es como se representa en el computador, es decir, tendremos la matriz (2D) de números como ya hemos mencionado anteriormente. Éstos son los valores que se manipulan para extraer información de las imágenes mediante programas (software).

En la siguiente figura se ilustra un ejemplo de la cuantización espacial y en amplitud:

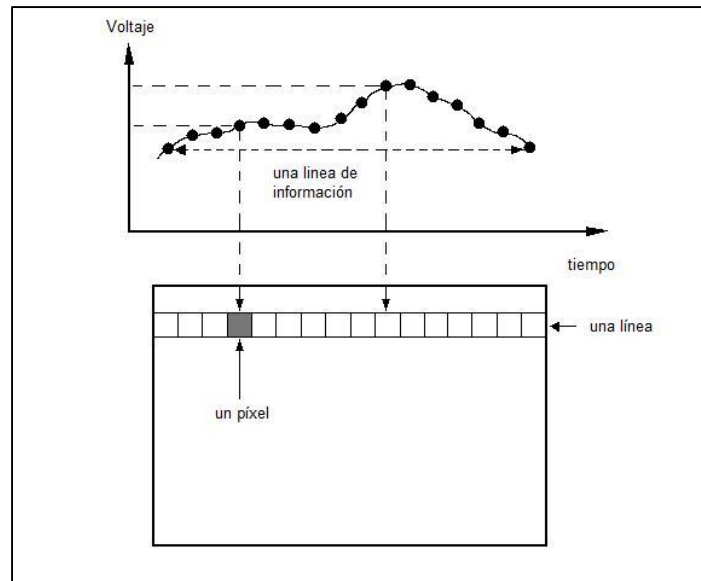


FIGURA 2: DIGITALIZACIÓN DE UNA SEÑAL ANALÓGICA

Supongamos que tenemos una señal analógica, que bien podría ser una línea de video analógica. Esta señal (línea) analógica de video se convierte a una imagen digital muestreando la señal analógica a intervalos determinados. El procedimiento consiste en medir el voltaje de la señal a intervalos de tiempo fijos. El valor del voltaje en cada instante se convierte a un número que es almacenado y se corresponde con la intensidad de la imagen en ese punto. La intensidad en cada punto depende tanto de las propiedades intrínsecas del objeto que se está viendo como de las condiciones de luz de la escena. Repitiendo este procedimiento para todas las líneas de video que constituyen una imagen, se pueden grabar los resultados obtenidos en el computador, de suerte que habremos conseguido una imagen digital que, en definitiva, es una matriz de números.

La imagen puede accederse como una matriz bidimensional (2D) de datos, donde cada punto o dato se denomina píxel.

Además de las cámaras de televisión que generan una imagen de video, uno de los sensores más usados para la visión artificial son los dispositivos de acoplamiento de carga (Charge Coupled Devices - CCD). Entre los dispositivos CCD, que generalmente también producen una señal continua de video, cabe distinguir dos categorías: sensores de exploración de línea y sensores de exploración de área. Estos sensores CCD se basan en unos elementos

semiconductores llamados photosites. Los fotones procedentes de la escena excitan el elemento semiconductor, de forma que el grado de excitación es proporcional a la cantidad de carga acumulada en el photosite y, por lo tanto, a la intensidad luminosa en ese punto.

Estos photosites se pueden representar en forma de matriz como se muestra en la siguiente figura:

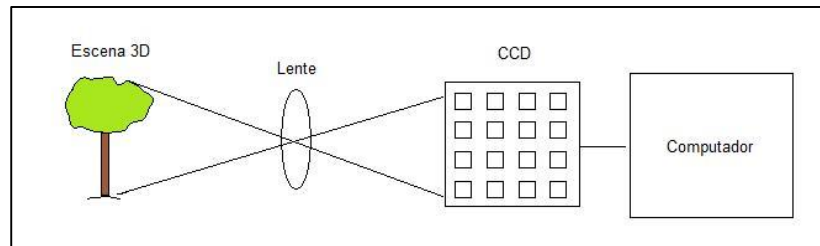


FIGURA 3: CAPTURA DE UNA IMAGEN 3D POR UN DISPOSITIVO CCD

Supongamos una matriz de photosites situados detrás de una lente y sobre los que se proyecta una imagen procedente de la escena 3D. La señal de estos sensores se procesa en el propio sensor (cámara) o en otro dispositivo (tarjeta de procesamiento de imágenes digitales) y los valores digitales se envían al computador.

Para clarificar un poco más estos conceptos, analicemos el ejemplo sencillo mostrado en la figura siguiente:

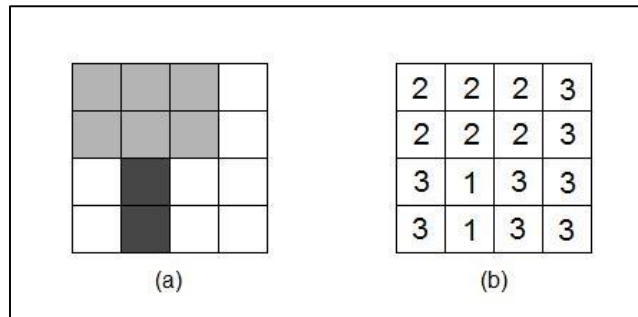


FIGURA 4: FIGURA DEL ÁRBOL CAPTURADA POR UNA CÁMARA CON 4X4 SENSORES DE INTENSIDAD

Supongamos una escena 3D cuya representación digital en forma de matriz, tal como se almacena en el computador, resulta ser la que se muestra en la figura 3-(b). En este caso se utiliza una matriz de 16 elementos (4x4) para representar el árbol; lógicamente, cuantos más sensores dispongamos en la cámara, más precisión podemos tener en la reproducción de la imagen. Esto es lo que se conoce como resolución espacial de la imagen.

2.3.9. IMÁGENES BLANCO/NEGRO Y COLOR

En definitiva, e independientemente del tipo de sensor utilizado, la imagen que ha de ser tratada por el computador se presenta digitalizada espacialmente en forma de matriz con una resolución de $M \times N$ elementos.

Si la imagen es en Blanco y Negro (B/N), se almacena un valor por cada píxel. Se suele utilizar un rango de valores para su representación, que generalmente es de 0 a $2^n - 1$. Uno de los valores más utilizados de n es 8; esto significa que el rango de valores para este caso varía de 0 a 255. En este caso, el 0 representa el negro absoluto y el 255, el blanco absoluto. Esto indica que podemos tener una resolución o precisión en los grises posibles de 256. El hecho de utilizar 256 niveles es porque con 8 bits del computador se pueden codificar 256 valores distintos desde la combinación 00000000, que representa el nivel 0, hasta la combinación 11111111, que representa el nivel 255.

En el caso de las imágenes en color, los elementos de la matriz vienen dados por tres valores, que representan cada uno de los componentes básicos del color en cuestión. Estos componentes son el Rojo (R), Verde (G) y Azul (B), el

conocido código RGB. En este caso el conjunto de valores (0,0,0) es el negro absoluto; el (255,255,255), el blanco absoluto; el (255,0,0), el rojo puro; el (0,255,0), el verde puro; el (0,0,255), el azul puro. Como es lógico, la combinación de distintos valores proporciona otros colores, por ejemplo, el (255, 255, 51) es un tono de amarillo o el (204, 153, 102), es un tono marrón. El número de colores posible resulta ser 255.

2.3.10. RESOLUCIÓN ESPACIAL Y EN AMPLITUD

Para comprender mejor el sentido y la diferencia que existe entre resolución espacial y en amplitud, vamos a ilustrar los conceptos en las siguientes figuras:

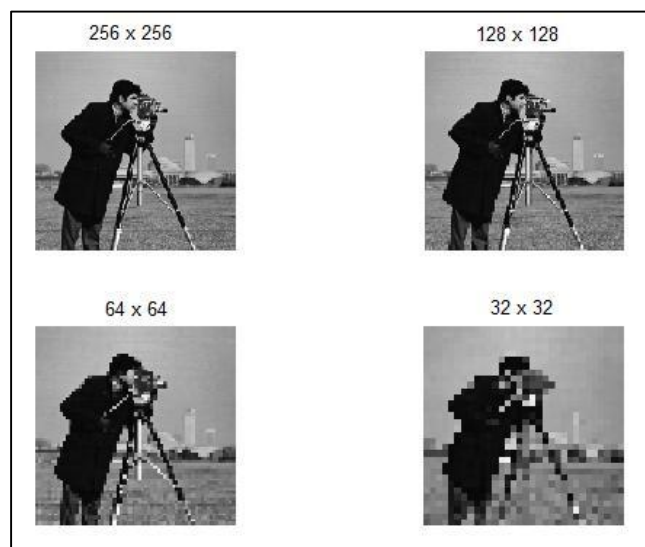


FIGURA 5: CUATRO REPRESENTACIONES DE LA MISMA IMAGEN CON VARIACIÓN EN EL NÚMERO DE PÍXELES UTILIZADOS



FIGURA 6: SEIS REPRESENTACIONES DE LA MISMA IMAGEN CON VARIACIÓN EN EL NÚMERO DE NIVELES DE GRIS UTILIZADOS

Como se muestra en las figuras anteriores, dependiendo del número de píxeles que tenga el dispositivo y de niveles de grises con que se trabaje en el computador, la imagen poseerá más o menos resolución espacial y en amplitud respectivamente.

2.3.11. REPRESENTACIÓN DE IMÁGENES DIGITALES

Como ya se ha mencionado antes, el termino imagen se refiere a una función de intensidad bidimensional, la cual puede ser representada como $f(x,y)$, $f(i,j)$, $I(x,y)$, $I(i,j)$, etc., donde x e y , o bien, i y j son las coordenadas espaciales y el valor de f o I en cualquier punto (x,y) o (i,j) es proporcional a la intensidad o nivel de gris de la imagen en ese punto.

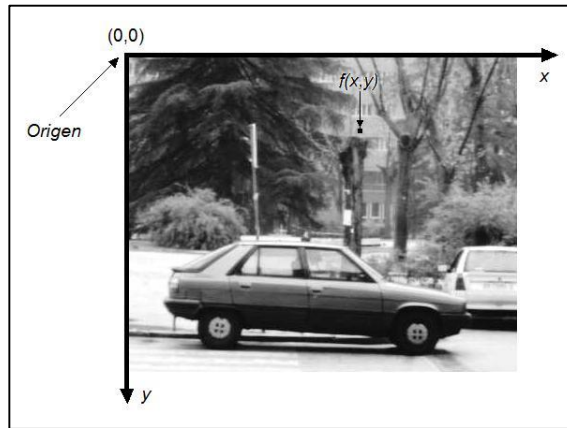


FIGURA 7: CONVENCION DE EJES UTILIZADA PARA LA REPRESENTACION DE IMÁGENES DIGITALES

2.4. PROCESAMIENTO Y ANALISIS DE IMÁGENES DIGITALES

Aunque la distinción entre procesamiento y análisis de imágenes digitales no es obvia de forma inmediata, el procesamiento de imágenes puede ser visto como una transformación de una imagen a otra imagen, es decir, a partir de una imagen, se obtiene otra imagen modificada. Por otro lado, el análisis es una transformación de una imagen en algo distinto a una imagen; en consecuencia, el análisis es un determinado tipo de información representando una descripción o una decisión. En la mayoría de los casos, las técnicas de análisis de imágenes digitales son aplicadas a imágenes que han sido procesadas previamente. Desde el punto de vista de un observador humano, el análisis de imágenes es una tarea fácil y rápida, algo que no ocurre en visión artificial. Igualmente, la capacidad de percepción humana permite procesar rápidamente una imagen para detectar en ella características de interés, por ejemplo, bordes o regiones.

2.5. PROCESAMIENTO BASICO DE IMÁGENES

El procesamiento de datos en el sistema de visión puede enfocarse desde 2 perspectivas:

Alteración píxel a píxel de los datos en una escala global (individuales).

Operaciones basadas en múltiples puntos (vecindad).

La generación de un nuevo píxel en una imagen será una función bien del valor de cada píxel en su localización individual, o bien de los valores de los píxeles en la vecindad de un píxel dado, como se indica en la figura siguiente:

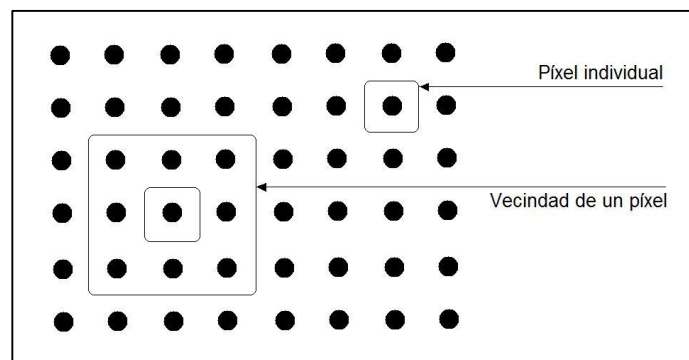


FIGURA 8: FUNCIONES DE PUNTO Y VECINDAD

Existen aún otras operaciones, que no se clasifican ni como individuales ni como vecindad ya que transforman las imágenes por otros procedimientos; son transformaciones que operan globalmente sobre los valores de intensidad de la imagen cuyo efecto es un realzado de la imagen original, operaciones aritméticas y lógicas, estas últimas basadas en la teoría del álgebra de Boole, y operaciones que realizan transformaciones geométricas, sin modificar los valores de intensidad.

2.5.1. OPERACIONES INDIVIDUALES

Las operaciones individuales implican la generación de una nueva imagen modificando el valor del píxel en una simple localización basándose en una regla global aplicada a cada localización de la imagen original. El proceso consiste en obtener el valor del píxel de una localización dada en la imagen, modificándolo por una operación lineal o no lineal y colocando el valor del nuevo píxel en la

correspondiente localización de la imagen nueva. El proceso se repite para todas y cada de las localizaciones de los píxeles en la imagen original.

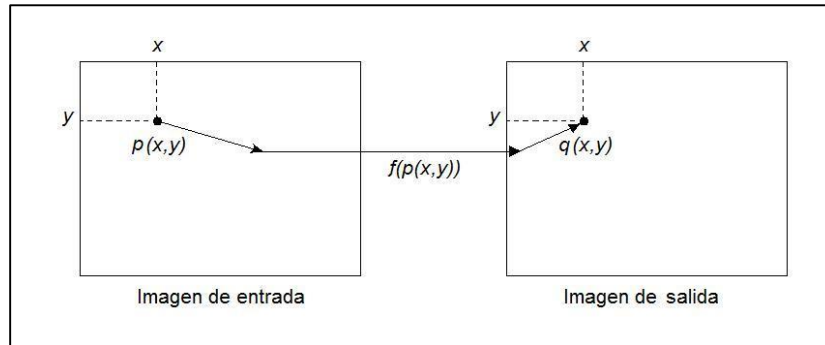


FIGURA 9: OPERACIÓN INDIVIDUAL

Como se aprecia en la figura anterior, el operador individual es una transformación uno a uno. El operador f se aplica a cada píxel en la imagen o sección de la imagen y la salida depende únicamente de la magnitud del correspondiente píxel de entrada; la salida es independiente de los píxeles adyacentes. La función transforma el valor del nivel de gris de cada píxel en la imagen y el nuevo valor se obtiene a través de la ecuación:

$$q(x,y) = f(p(x,y))$$

La función f puede ser un operador lineal o no lineal. El proceso matemático es relativamente simple. La imagen resultante es de la misma dimensión que la original.

2.5.2. OPERADOR IDENTIDAD

Este operador crea una imagen de salida que es idéntica a la imagen de entrada. La función de transformación es:

$$q=p$$

El operador identidad deja la imagen de entrada invariante. En la siguiente figura se muestra la función de transformación dada por la ecuación anterior:

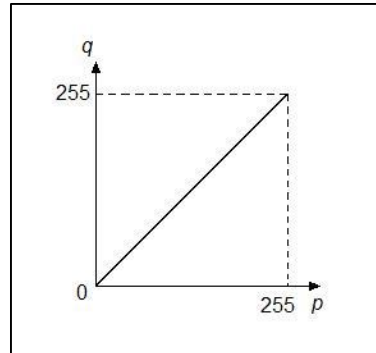


FIGURA 10: REPRESENTACIÓN DEL OPERADOR IDENTIDAD

2.5.3. OPERADOR INVERSO O NEGATIVO

Este operador crea una imagen de salida que es inversa de la imagen de entrada. Este operador es útil en diversas aplicaciones tales como imágenes médicas. Para una imagen con valores de gris en el rango de 0 a 255 la función de transformación resulta ser:

$$q = 255 - p$$

En la siguiente figura se muestra la función de transformación dada por la ecuación anterior:

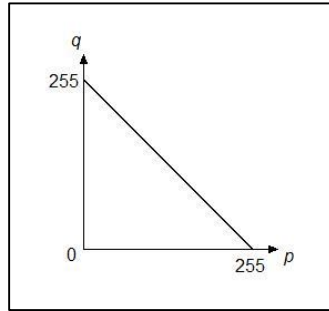


FIGURA 11: REPRESENTACIÓN DEL OPERADOR INVERSO

2.5.4. OPERADOR UMBRAL

Esta clase de transformación crea una imagen de salida binaria a partir de una imagen de grises, donde el nivel de transición está dado por el parámetro de entrada p_1 . La función de transformación es la siguiente:

$$q = \begin{cases} 0 & \text{para } p \leq p_1 \\ 255 & \text{para } p > p_1 \end{cases}$$

En la siguiente figura se muestra la función de transformación dada por la ecuación anterior:

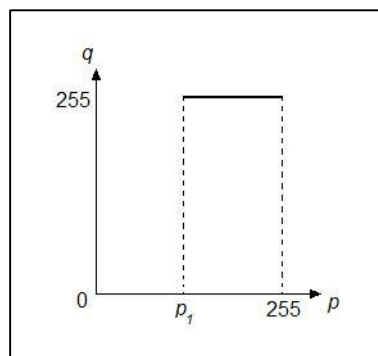


FIGURA 12: REPRESENTACIÓN DEL OPERADOR UMBRAL

2.5.5. TRANSFORMACIÓN DE VECINDAD

En las operaciones de vecindad modifican el nuevo valor del píxel en la imagen de salida depende de una combinación de los valores de los píxeles en la vecindad de la imagen original que está siendo transformada.

Dentro de la categoría de operaciones de vecindad se incluyen las operaciones de filtrado. Las operaciones de filtrado tienen la particularidad de eliminar un determinado rango de frecuencias de las imágenes.

2.5.5.1. Nociones y propiedades de vecindad

Se dice que todo píxel p , de coordenadas (x,y) , tiene cuatro píxeles que establecen con él una relación de vecindad horizontal o vertical, que son:

Horizontal: $(x-1, y)$ y $(x+1, y)$ Vertical: $(x, y-1)$ y $(x, y+1)$

Estos cuatro píxeles definen lo que se conoce como entorno de vecindad-4 y nos referimos a ellos como $E_4(p)$.

Los cuatro vecinos diagonales de p tienen coordenadas:

$(x-1, y-1), (x+1, y-1), (x-1, y+1), (x+1, y+1)$

y nos referimos a ellos como $E_D(p)$. Estos píxeles junto con los $E_4(p)$ se llaman los vecinos-8 de p y se denotan como $E_8(p)$.

Existen excepciones dadas cuando el píxel (x, y) es un punto del borde de la imagen, en cuyo caso algunos de los vecinos definidos anteriormente no existen.

2.5.5.2. Conectividad

Sea V el conjunto de valores de intensidad de los píxeles que se permiten estén adyacentes, por ejemplo, si sólo se desea que exista conectividad entre los píxeles

con intensidades 80,81 y 83, entonces $V = \{80,81,83\}$. Consideremos tres tipos básicos de conectividad:

Conectividad-4. Dos píxeles p y q con valores de V están 4-conectados si q está en el conjunto $E_4(p)$.

Conectividad-8. Dos píxeles p y q con valores de V están 8-conectados si q está en el conjunto $E_8(p)$.

Conectividad-m (mixta). Dos píxeles p y q con valores de V están m-conectados si q está en el conjunto $E_4(p)$ o q está en $E_D(p)$ y $E_4(p) \cap E_4(q) = \emptyset$.

Un píxel p es contiguo a otro píxel q si están conectados. Se puede definir la adyacencia-4, 8 o m, dependiendo del tipo de conectividad especificada. Dos subconjuntos imagen $S1$ y $S2$ son contiguos si algún píxel de $S1$ es contiguo a algún píxel de $S2$.

Un camino desde el píxel p con coordenadas (x, y) hasta un píxel q con coordenadas (s, t) es una secuencia de varios píxeles con coordenadas,

$$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$$

donde $(x_0, y_0) = (x, y)$ y $(s, t) = (x_n, y_n)$ son adyacentes a (x_{i-1}, y_{i-1}) , con $1 \leq i \leq n$, y n es la longitud del camino. Se pueden definir caminos -4, 8 o m dependiendo del tipo de adyacencia usada.

Si p y q son píxeles de un subconjunto de imagen S , entonces p está conectado a q en S si existe un camino desde p hasta q formado de píxeles pertenecientes a S . Dado un píxel p cualquiera de S , el conjunto de píxeles de S que están conectados a p se llaman componente conectado de S . Se deduce que dos píxeles cualesquiera de un componente conectado están a su vez conectados entre sí y que los componentes conectados distintos son disjuntos.

Un camino simple es un camino sin píxeles repetidos y un camino cerrado es un camino simple en el cual el primer píxel es un vecino del último.

2.6. OPERACIONES DE FILTRADO

Las operaciones de filtrado basan su operatividad en la Convolución de la imagen utilizando el denominado núcleo de Convolución. Cabe distinguir dos tipos de filtros: paso alto y paso bajo, que en el contexto de la teoría de señales supone que los primeros dejan pasar las altas frecuencias de la señal y los segundos, las bajas. En el caso de las imágenes nos referimos a frecuencias espaciales. De una forma, las altas frecuencias se asocian a cambios bruscos de intensidad en pequeños intervalos espaciales, es decir, bordes, mientras que las bajas frecuencias se refieren a cambios lentos en la intensidad.

2.6.1. FILTROS PASA BAJO

En la siguiente figura se muestran algunos núcleos de convolución que caracterizan los filtros paso bajo:

$$PB_1 \equiv \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$PB_2 \equiv \frac{1}{10} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$PB_3 \equiv \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

FIGURA 13: TRES NÚCLEOS REPRESENTATIVOS DE FILTROS PASO BAJO

2.6.2. FILTROS PASA ALTO

En la siguiente figura se muestran algunos núcleos de Convolución que caracterizan los filtros paso alto:

$$PA_1 \equiv \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

$$PA_2 \equiv \begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

$$PA_3 \equiv \begin{bmatrix} 1 & -2 & 1 \\ -2 & 5 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

FIGURA 14: TRES NÚCLEOS REPRESENTATIVOS DE FILTROS PASO ALTO

2.7. HISTOGRAMA

El histograma de una imagen es la representación gráfica o analítica de la distribución relativa de cada valor posible de pixel de imagen, y en caso de imágenes grises de 8 bits será un vector de 256 componentes, siendo la componente i el número de pixeles de nivel i en la imagen, dividido por el número total de pixeles:

El histograma es formalmente la función estadística de densidad de probabilidad en forma discreta de los distintos niveles de gris dentro de la imagen.

$$\text{histograma}[i] = \frac{N_{\text{pixeles de nivel } i}}{N_{\text{total}}} \quad 0 \leq i \leq 255$$

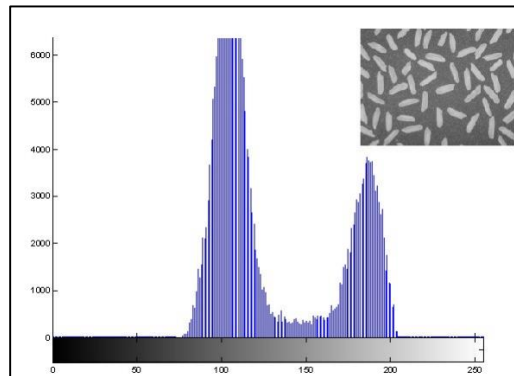


FIGURA 15: HISTOGRAMA

2.8. ESPACIOS DE COLORES Y COLORIMETRÍA

Como ya se mencionó en una sección anterior, es posible representar los colores en base a una composición de colores fundamentales, imitando lo que hace nuestro ojo. En base a este hecho, es que se han creado espacios de colores que de alguna manera u otra, permiten representar un conjunto de valores. Normalmente no se pueden representar todos los colores que son realmente visibles, sin embargo, son lo suficiente como para representar la gran mayoría de colores que vemos regularmente.

Un espacio de color define un modelo de composición del color. Por lo general un espacio de color lo define una base de N vectores (por ejemplo, el espacio RGB lo forman 3 vectores: Rojo, Verde y Azul), cuya combinación lineal genera todo el espacio de color. Los espacios de color más generales intentan englobar la mayor cantidad posible de los colores visibles por el ojo humano, aunque existen espacios de color que intentan aislar tan solo un subconjunto de ellos.

Existen espacios de color de:

Una dimensión: escala de grises, escala Jet, etc.

Dos dimensiones: sub-espacio rg, sub-espacio xy, etc.

Tres dimensiones: espacio RGB, HSV, YCbCr, YUV, Y'I'Q', etc.

Cuatro dimensiones: espacio CMYK.

De los cuales, los espacios de color de tres dimensiones son los más extendidos y los más utilizados para la mayoría de las aplicaciones. En el caso que se cuente con un espacio de una dimensión de profundidad, estaríamos en presencia de un espacio de color que en la práctica puede representar el grado de luminosidad de la imagen; es básicamente una imagen en escala de grises, y no entrega información de color. Pero si queremos trabajar con colores, los modelos existentes especifican tres coordenadas, o atributos, que representan su posición dentro de un espacio de color específico. Estas coordenadas no nos dicen cuál es el color, sino que muestran dónde se encuentra un color dentro de un espacio de color en particular. A continuación, se hará una descripción de los espacios de color más relevantes, así como los más importantes para este trabajo.

2.8.1. ESPACIO DE COLOR RGB

Este modelo es el más utilizado para visualizar imágenes digitales en una pantalla en los formatos actuales. Como se ha dicho anteriormente, es un modelo aditivo, en el que sumando distintas cantidades de colores primarios podemos conseguir otros colores. En los archivos gráficos que utilizan este modelo utilizamos para representar una tripleta (R, G, B).

Contienen tres planos de imágenes independientes, uno por cada color primario. Cuando estos planos se inyectan a un monitor que utiliza este sistema, la pantalla de fósforo reproduce una imagen a color. Es un sistema aditivo que, variando la cantidad de color rojo, verde y azul agregados al negro, se produce nuevos colores. En los archivos gráficos el sistema RGB se usa para representar cada píxel con una tripleta o terna de la forma (R, G, B), quedando representado en un sistema cartesiano, como se puede apreciar en la figura.

En este sistema los colores vienen definidos por un punto del cubo, así los vértices comunes son los colores secundarios, el origen de coordenadas es el negro, y el punto donde se sumarian los 3 colores primarios será el vértice del color blanco. Y si trazamos una línea que una el negro con el blanco representará toda la escala de grises.

A un píxel se le asigna un valor entre 0(negro) y 255(blanco). La idea es la siguiente: si por ejemplo queremos un color rojo tendríamos que darle un valor alto a la coordenada del rojo y valores más pequeños a las otras dos coordenadas. Si todas las coordenadas son 255 tendremos un blanco puro e igualmente si todas son 0 tendremos un negro. Y si las 3 son iguales entonces tendremos un gris.

Las imágenes que utilizan este modelo pueden reproducir en la pantalla hasta 16.7 millones de colores, ya que tendremos 3 canales (que presentan a cada uno de los planos) y 8 bits para presentar 255 valores, es decir tendremos 3x8 bits de información de color para cada píxel.

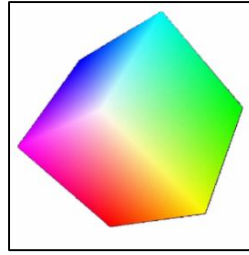


FIGURA 16: MODELO ADITIVO RGB

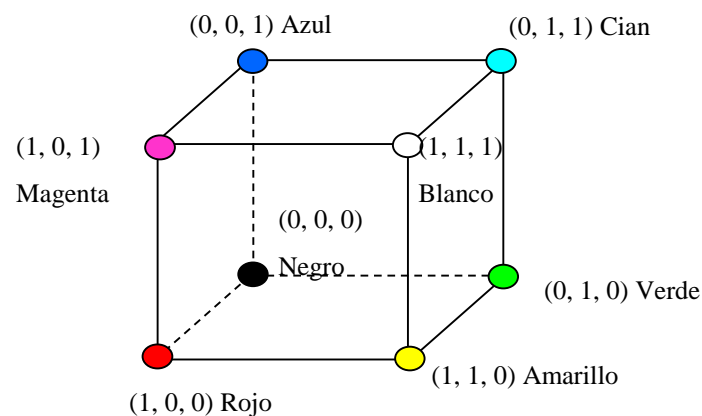


FIGURA 17: PESO DE LAS 3 COMPONENTES EN EL MODELO RGB

2.8.2. ESPACIO DE COLOR HSV

El modelo HSV (del inglés Hue, Saturation, Value) (Tonalidad, Saturación, Valor), también llamado HSB (Hue, Saturation, Brightness) (Tonalidad, Saturación, Brillo), define un modelo de color en términos de sus componentes constituyentes en coordenadas cilíndricas:

Tonalidad, el tipo de color (como rojo, azul o amarillo). Se representa como un grado de ángulo cuyos valores posibles van de 0 a 360° (aunque para algunas aplicaciones se normalizan del 0 al 100 %). Cada valor corresponde a un color. Ejemplos: 0 es rojo, 60 es amarillo y 120 es verde.

Saturación. Se representa como la distancia al eje de brillo negro-blanco.

Los valores posibles van del 0 al 100 %. A este parámetro también se le suele llamar "pureza" por la analogía con la pureza de excitación y la pureza colorimétrica de la colorimetría. Cuanto menor sea la saturación de un color, mayor tonalidad grisácea habrá y más decolorado estará. Por eso es útil definir la insaturación como la inversa cualitativa de la saturación.

Valor del color, el brillo del color. Representa la altura en el eje blanco negro.

Los valores posibles van del 0 al 100 %. 0 siempre es negro. Dependiendo de la saturación, 100 podría ser blanco o un color más o menos saturado.

La figura muestra una representación del espacio HSV.

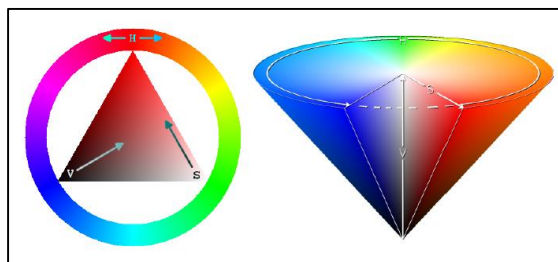


FIGURA 18: MODELO HSV

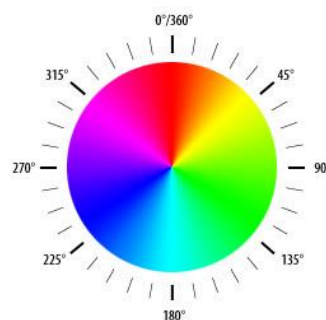


FIGURA 19: TINTE



FIGURA 20: SATURACIÓN



FIGURA 21: VALOR

2.9. TESSERACT OCR⁵

Tesseract es un motor OCR libre. Fue desarrollado originalmente por Hewlett Packard como software propietario entre 1985 y 1995. Tras diez años sin ningún desarrollo, fue liberado como código abierto en el año 2005 por Hewlett Packard y la Universidad de Nevada, Las Vegas. Tesseract es desarrollado actualmente por Google y distribuido bajo la licencia Apache, versión 2.0.

2.9.1. HISTORIA

El motor Tesseract se desarrolló en los laboratorios de Bristol de Hewlett Packard en Greeley (Colorado) entre 1985 y 1994. En 1996 se realizaron las modificaciones necesarias para su portabilidad en Windows, y más tarde, en 1998 se migró el sistema de C a C++.

2.9.2. CARACTERÍSTICAS

En 1995, Tesseract era uno de los tres mejores motores OCR en cuanto a precisión, además está disponible para Linux, Windows y Mac OS X, sin embargo, sólo ha sido probado por los desarrolladores en Windows y Ubuntu. Hasta la versión 2, Tesseract sólo podía aceptar como entrada imágenes de una sola columna en formato TIFF. En estas primeras versiones no se incluía análisis de patrones, y por tanto, las imágenes con múltiples columnas o anotaciones producían resultados ilegibles. Desde la versión 3, Tesseract soporta el formato en el texto y el análisis del patrón de la página. A través de la biblioteca Leptonica, se consigue la compatibilidad con nuevos formatos de imagen, además, se puede detectar si el texto proporcional o monoespaciado. Tesseract puede procesar inglés, francés, italiano, alemán, español, portugués brasileño y neerlandés, y puede ser entrenado para funcionar con otros idiomas.

Se puede encontrar en la siguiente página: <https://github.com/tesseract-ocr/>

2.10. GLOSARIO DE TÉRMINOS BÁSICOS

SISTEMA: Un sistema es un conjunto ordenado de elementos que se encuentran insurreccionados entre sí para lograr un fin en común.

⁵ https://es.wikipedia.org/wiki/Tesseract_OCR

SIMULACIÓN: Simulación es una técnica numérica para conducir experimentos en una computadora digital. Estos experimentos comprenden ciertos tipos de relaciones matemáticas y lógicas, las cuales son necesarias para describir el comportamiento y la estructura de sistemas complejos del mundo real a través de largos períodos.

BLOB ANÁLISIS: Un algoritmo utilizado en visión artificial que identifica objetos segmentados y los mide según parámetros morfométricos (tamaño, diámetro, perímetro, etc.) o densitométricos (nivel de gris medio, media, color)

ATRIBUTO: Un atributo es una especificación que define una propiedad de un Objeto, elemento o archivo. También puede referirse o establecer el valor específico para una instancia determinada de los mismos.

CCD: es la abreviatura de Charge-Coupled Device. Un sensor CCD es un dispositivo semiconductor sensible a la luz, que convierte las partículas de luz (fotones) en cargas eléctricas (electrones). Las cámaras CCD son una de los dos tipos de cámaras que dominan el mercado de la visión, conjuntamente con las cámaras CMOS.

CMOS: Es el acrónimo de Complementary Metal Oxide Semiconductor. Esta tecnología funciona como un fotodiodo donde la luz genera una corriente que es representativo de la cantidad de luz que impacta en cada píxel, por tanto difiere significativamente de la tecnología CCD. Hay un número de ventajas en la utilización de los sensores CMOS con respecto a los CCD entre los que se destaca el coste, velocidad, anti-blooming, y características de respuesta programable (Ej. Respuesta con múltiples pendientes).

ESCALA DE GRISES: En visión este término se refiere a una imagen monocroma con una gradación de grises. Una cámara monocroma que trabaje a 8 bits generará una imagen con 256 tonos de gris. Si la cámara trabaja a 12 bits su escala de grises será de 4096 niveles.

HISTOGRAMA: Es una representación gráfica de los valores de todos los píxeles de la imagen. Generalmente a la izquierda de la gráfica se representa el valor cero correspondiente al negro y a la derecha se representa el valor máximo (que depende del número de bits) correspondiente al blanco. La curva del histograma representa la cantidad de píxeles de cada valor de gris que hay en la imagen. Si la

imagen es en color se acostumbran a representar los histogramas de valores de rojo, verde y azul.

MODELO: Un modelo es una representación de un objeto, sistema o idea, de forma diferente al de la entidad misma. El propósito de los modelos es ayudarnos a explicar, entender o mejorar un sistema. Un modelo de un objeto puede ser una réplica exacta de éste o una abstracción de las propiedades dominantes del objeto.

ENTIDAD: El término entidad o ente, en su sentido más general, se emplea para denominar todo aquello cuya existencia es perceptible por algún sistema animado, véase; lógica o semántica. Una entidad puede por lo tanto ser concreta, abstracta, particular o universal. Es decir, las entidades no son sólo los objetos cotidianos como sillas o personas, sino también propiedades, las relaciones, los eventos, números, conjuntos, proposiciones, mundos posibles, creencias, pensamientos, y otros.

SISTEMA EXPERTO: Los Sistemas Expertos, rama de la Inteligencia Artificial, son sistemas informáticos que simulan el proceso de aprendizaje, de memorización, de razonamiento, de comunicación y de acción en consecuencia de un experto humano en cualquier rama de la ciencia.

INTELIGENCIA ARTIFICIAL: Es aquella inteligencia exhibida por artefactos creados por humanos (es decir, artificial). A menudo se aplica hipotéticamente a los computadores. El nombre también se usa para referirse al campo de la investigación científica que intenta acercarse a la creación de tales sistemas.

ROBOTICA: Es la rama de la tecnología que se dedica al diseño, construcción, operación, disposición estructural, manufactura y aplicación de los robots. La robótica combina diversas disciplinas como son: la mecánica, la electrónica, la informática, la inteligencia artificial, la ingeniería de control y la física. Otras áreas importantes en robótica son el álgebra, los autómatas programables, la animatrónica y las máquinas de estados.

DINAMICA DE SISTEMAS: Dinámica de Sistemas es un nombre propio que designa un determinado método de construcción de modelos de sistemas sociales susceptibles de ser simulados por ordenador. El origen de esta técnica se remonta a finales de los años cincuenta y su implantación definitiva se produce durante la década de los sesenta. El desarrollo de este método se debe al trabajo de J. W.

Forrester del Instituto Tecnológico de Massachussets, el cual por primera vez utilizó técnicas pertenecientes a las disciplinas de ingeniería automática para el estudio de procesos sociales y económicos.

VISION ARTIFICIAL: La visión artificial, también conocida como visión por computador (del inglés computer vision) o visión técnica, es un subcampo de la inteligencia artificial. El propósito de la visión artificial es programar un computador para que "entienda" una escena o las características de una imagen. (La detección, segmentación, localización y reconocimiento de ciertos objetos en imágenes, por ejemplo, caras humanas).

REDES NEURONALES: Las redes neuronales son sistemas ideados como abstracciones de las estructuras neurobiológicas (cerebros) encontradas en la naturaleza y tienen la característica de ser sistemas desordenados capaces de guardar información. La forma en que desarrollan su trabajo es esencialmente distinta de la utilizada por las computadoras convencionales. Los procesadores microscópicos del cerebro (neuronas) operan en paralelo y presentan cualitativamente más ruido que los elementos que forman a las computadoras. No ejecutan un programa fijo con base en un conjunto previamente especificado de datos, sino que comunican señales a través de retransmisores que llamamos sinapsis.

LOGICA DIFUSA: La lógica difusa es una metodología que proporciona una manera simple y elegante de obtener una conclusión a partir de información de entrada vaga, ambigua, imprecisa, con ruido o incompleta, en general la lógica difusa imita como una persona toma decisiones basada en información con las características mencionadas. Una de las ventajas de la lógica difusa es la posibilidad de implementar sistemas basados en ella tanto en hardware como en software o en combinación de ambos.

2.11. HIPÓTESIS

Utilizando visión artificial, se reconocerá automáticamente el número del documento nacional de identificación de una persona.

III. MARCO METODOLÓGICO

3.1. MÉTODOS Y PROCEDIMIENTOS

Los materiales o equipos que se han utilizado para esta investigación son una computadora personal y una cámara webcam.

- Laptop de 2 GHz, 1GB RAM, 350GB DD.
- Una cámara webcam de 25 cuadros por segundo, resolución de 320 x 240 píxeles.

Para llevar a cabo la captura y el procesamiento digital adecuado, se hace necesario tener una iluminación adecuada, si es de día, la luz natural es suficiente de lo contrario si es de noche, basta con la iluminación del ambiente o también, se puede hacer uso de iluminación led. Está demás decir, que, con una buena iluminación, mejor captura de la escena, por lo tanto, mejor funcionarán los algoritmos, y habrá un buen reconocimiento.

El software utilizado, es el lenguaje de programación Visual C++, junto con las librerías de Visión Artificial, OpenCV. Se escogió este lenguaje, por la facilidad y la velocidad de procesamiento rápido. El sistema operativo de la PC, sobre los que se ejecutan los programas, es Windows 8.1.

Para el desarrollo de la investigación, lectura automática del DNI, se han seguido los siguientes pasos:

3.1.1. CAPTURA

El tamaño de la captura de la imagen del DNI es una sola, por cuanto, la distancia de la cámara a la escena es fija. Esta distancia es de 15cm.

No se está considerando la rotación, por lo tanto, el DNI tendrá que ponerse o ubicarse de forma correcta.

Para capturar y ayudar la segmentación del DNI, de preferencia se busca poner el DNI en un fondo de color solido diferente del color celeste del DNI, que puede ser blanco o negro u otro diferente.

3.1.2. RGB A HSV

Dado que el color del DNI es de color celeste, la imagen se pasa del modelo RGB a HSV, para poder separar el objeto de interés del resto de la escena.

3.1.3. DEFINIR EL RANGO DEL COLOR CELESTE

El rango del color celeste es:

Hue: 180°

Saturación: 100%

Value: 100%

3.1.4. BINARIZAR DNI

Teniendo el color celeste, como único objeto de la escena, se Binariza. El resultado de este paso es una imagen binaria, con el objeto de color blanco y el fondo de color negro.

3.1.5. ETIQUETAR OBJETOS

Con los pasos anteriores, aparecerán objetos de diferentes tamaños. Para el siguiente paso, hay que etiquetarlos, así separamos todos los diferentes objetos que se encuentren en la escena.

3.1.6. SEGMENTAR EN BASE AL AREA

Si tenemos etiquetados los objetos, podemos saber el área de cada uno de ellos. Este parámetro servirá, para saber el objeto más grande, que es al que pertenece el objeto del DNI.

El valor del área mínimo es de 10000 píxeles. Se considera objeto del DNI, al objeto que tenga un área mayor a este valor y se obtiene una nueva imagen, sólo del DNI.

3.1.7. DIVIDIR EN 3 REGIONES HORIZONTALES EL DNI

Ahora que tenemos solamente la imagen del DNI, podemos dividirla e 3 partes horizontales, la cual, la tercera región inferior, será la que contiene el número del DNI. Hasta aquí, el procesamiento anterior nos ha servido solamente para segmentar la región donde se encuentra el número del DNI.

3.1.8. BINARIZACIÓN DE 3ª REGIÓN A PARTIR DE LA IMAGEN ORIGINAL

Se hace uso del algoritmo de binarización, con un valor de umbral de 100 para que aparezcan los caracteres en color blanco y el fondo de color negro. Este resultado se obtiene a partir de la imagen en escala de grises, que a su vez se obtiene de la imagen en color RGB.

3.1.9. ETIQUETAR OBJETOS EN LA 3ERA REGION DE INTERÉS

Como hemos dividido la imagen del DNI, el siguiente paso es etiquetar los objetos que se encuentran en esta región, que pertenecerían a los caracteres.

3.1.10. BUSCAR EL CARÁCTER

Se comienza a buscar desde arriba hacia abajo y de izquierda a derecha de la imagen, el carácter R. Si se encuentra, los siguientes 8 caracteres pertenecerán al número del documento nacional de identidad, DNI.

3.1.11. LECTURA DEL DNI

La Clase TesseractOCR, es una Clase importada de un repositorio Github, contiene todos los métodos necesarios para la lectura de caracteres de una imagen optimizados para OpenCV.

IV. RESULTADOS Y DISCUSIÓN

4.1. RESULTADOS

Observamos la imagen de entrada, capturada a 15 cm con respecto a la cámara. La escena es capturada con iluminación natural, o sea de día. Se observa también que el DNI es físico sobre un fondo de color diferente pero sólido. También se observa sobre el fondo una raya negra vertical que no perjudica después la lectura.

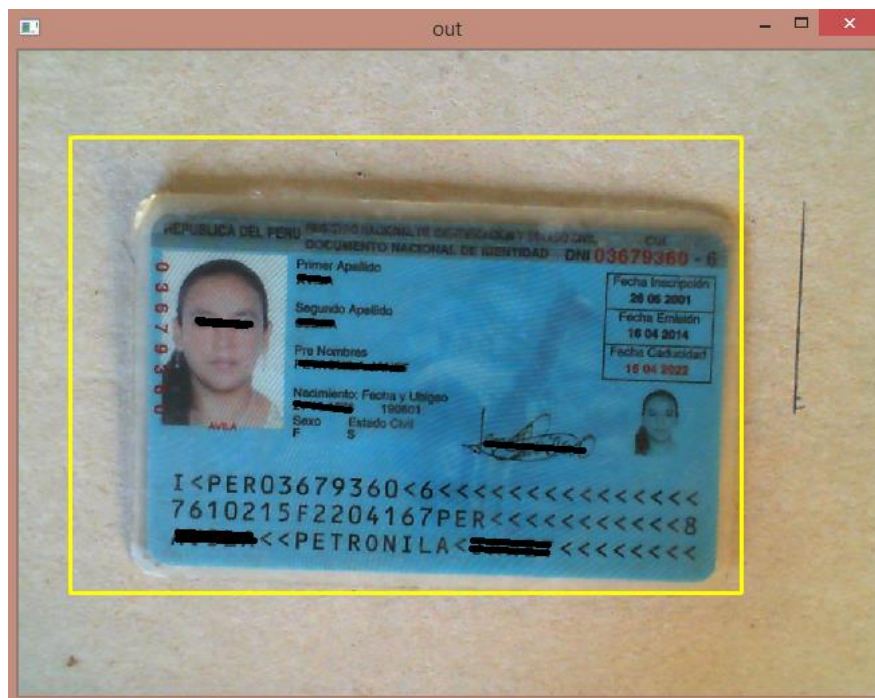
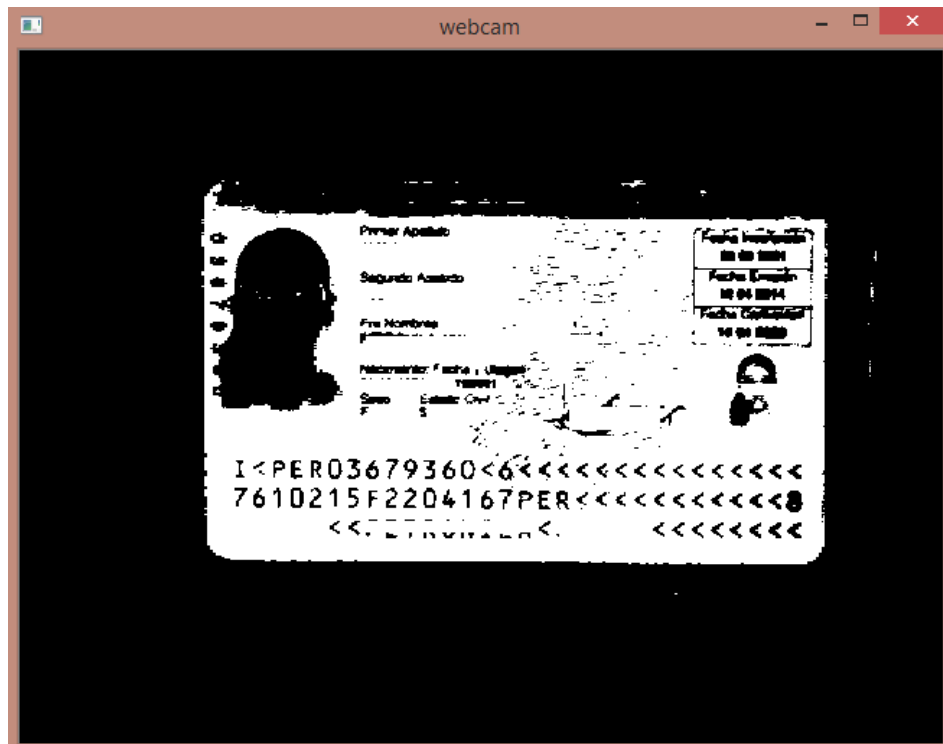


FIGURA 22: IMAGEN DE ENTRADA CON DNI

En la siguiente figura, se observa la segmentación del DNI, en base al color y el tamaño. Objetos que sean mayores a un valor de 10000 pixeles puede ser considerado objeto DNI.



A continuación, tenemos solamente la sub imagen del DNI. A partir de aquí se Binariza, y se obtiene lo siguiente.

La región de interés donde se encuentra el número, se hace un recorrido de arriba hacia abajo y de izquierda a derecha para encontrar la letra R, que será

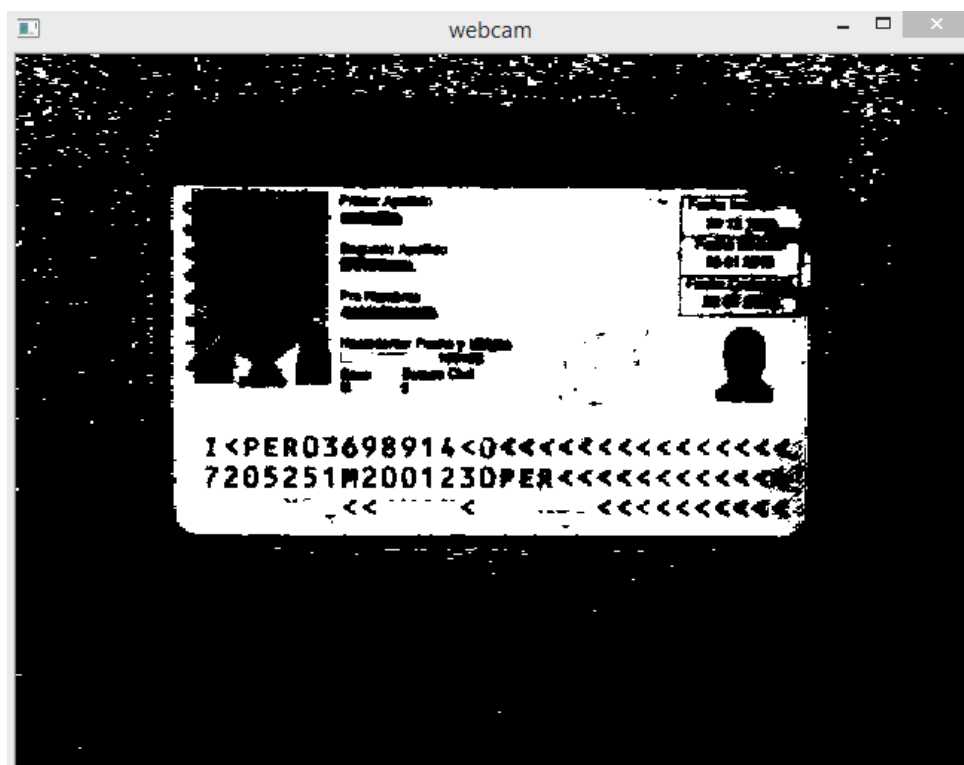
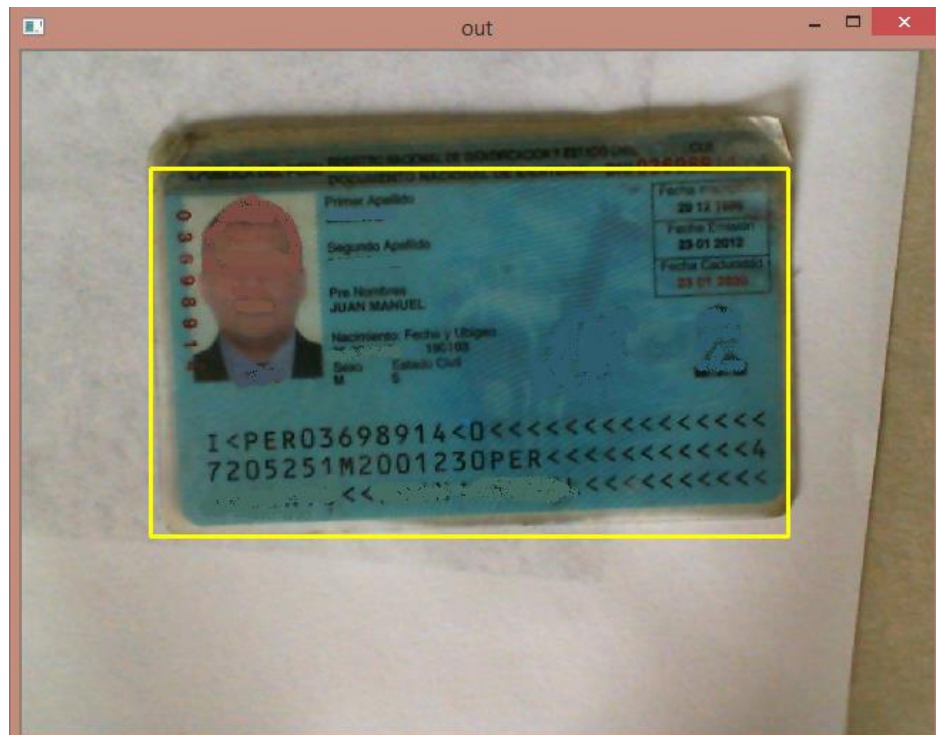




FIGURA 29: IMAGEN BINARIZADA

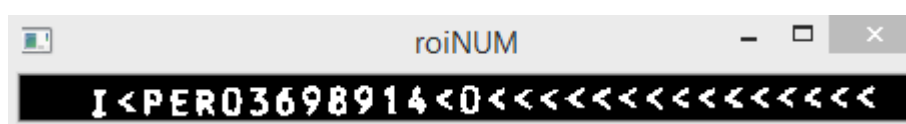
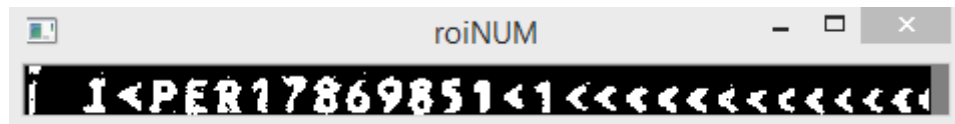


FIGURA 30: REGION DONDE SE ENCUENTRA EL NÚMERO



FIGURA 31: IMAGEN CON EL RESULTADO DE LA LECTURA





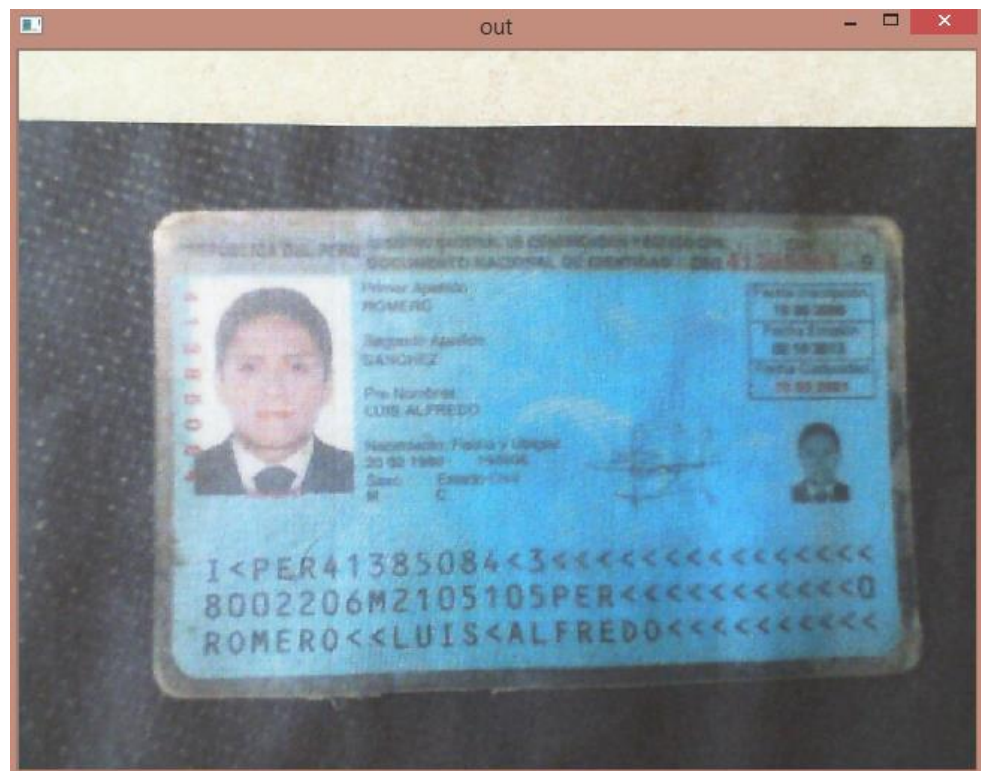
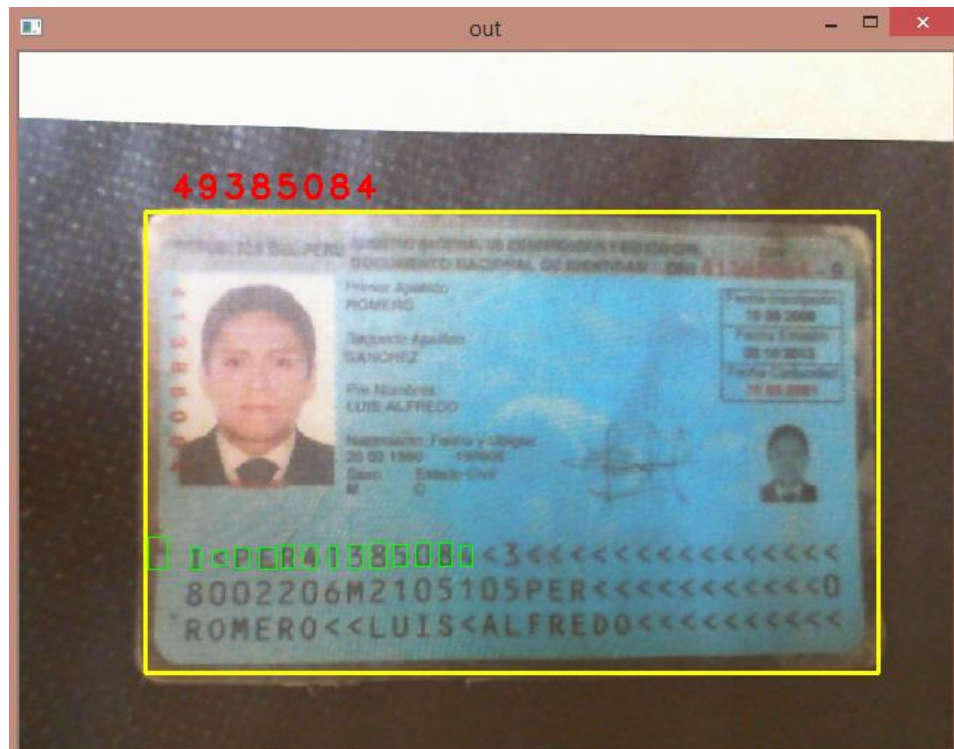


FIGURA 36: IMAGEN DE ENTRADA CON DNI



FIGURA 37: IMAGEN BINARIZADA



4.2. DISCUSIÓN

Como se puede observar, las dos identidades primeras se han procesado directamente con el DNI físico y las dos siguientes se han impreso en papel y sobre estas se ha hecho la captura y el procesamiento. Para las dos primeras, hay una mejor captura de las imágenes, mayor nitidez en comparación a las dos segundas. Se nota que hay mejor procesamiento en cuanto a las siguientes etapas, por lo tanto, menor probabilidad de que haya error en la etapa de la lectura.

También se observa diferentes iluminaciones, en algunas una iluminación uniforme y en otras una iluminación no adecuada. También se observa, la

captura no perpendicular con respecto a la cámara. Aun así, los algoritmos resultaron ser eficientes.

Se observa que el 100% de las imágenes capturadas son reconocidas. Esto se puede observar también, en el conjunto de imágenes bajadas de internet en el anexo N° 2. A pesar de haberse impreso (impresora de inyección de tinta) hay 100% de reconocimiento. Esto se debe a que los números no son de diferente tipo.

Cabe decir que la hipótesis planteada es verdadera, por cuanto es posible determinar el número del documento nacional de identidad, a través de visión artificial.

Esto sería un aporte importante, por lo que además del número del DNI, podríamos extraer otros datos de interés de este documento.

CONCLUSIONES

Se logró implementar un sistema de reconocimiento del número del Documento Nacional de Identidad, utilizando visión artificial.

Para la captura se tuvo en cuenta la iluminación y la distancia de la cámara a la escena, esto es importante, porque se obvia utilizar por ejemplo iluminación infrarroja y si la distancia no es fija, al momento de la captura se tiene diferentes tamaños del DNI, lo cual implicaría mas algoritmos de procesamiento.

Los DNI de color celeste se lograron mejorar, después de la captura, a través de un filtro suavizante. La segmentación de este se pudo lograr pasando del modelo RGB al modelo HSV, lo cual facilitó separar el DNI del fondo de la imagen.

El reconocimiento de los números se logró con las librerías de Reconocimiento Óptico de Caracteres, logrando una eficiencia muy alta, debido a que la imagen del DNI tiene buena resolución y la distancia de captura es a 15cm de la escena.

Se logró tener buenos resultados con una cámara webcam de medianas prestaciones en cuanto a resolución.

RECOMENDACIONES

Se recomienda que el sistema sea implementado en un celular Smart. Además, tener en cuenta la rotación y la distancia del DNI, que podría ser variable, así como la iluminación, que podría hacerse con el flash del celular.

Si se quiere incluir los DNI amarillos, habría que tener en cuenta los algoritmos de segmentación de color.

REFERENCIAS BIBLIOGRÁFICAS

1. Rafael C. Gonzales, Richard E. Woods, Tratamiento Digital de Imágenes. USA, Addison Wesley Iberoamericana, S. A. 1996. 755 pp.
2. Michael Alder, An Introduccion to Pattern Recognition, Mike Alder, 2001.
3. Abraham Kandel, Horst Bunke, Mark Last, Applied graph theory in computer visión and pattern recognition, Springer-Verlag Berlin Heidelberg, 2007.
4. Bernd Jahne, Digital image processing, Springer-Verlag Berlin Heidelberg, 2002.
5. Gonzalo Pajares Martinsanz, Jesús Manuel de la Cruz García, José Manuel molina pascual, Juan Cuadrado Pardo, Alejandro López Correa, “Imágenes digitales - Procesamiento practico con java, Alfa Omega Grupo Editor S.A. 2004
6. Elaine Rich, Kevin Knight, “Inteligencia Artificial”, McGraw-Hill / Interamericana de España S.A. 1994

ANEXOS

ANEXO 1: CÓDIGO EN C++ DEL PROGRAMA

```
#include <opencv2\opencv.hpp>
#include "Tesseract.h"

using namespace cv;
using namespace std;
Mat mejorarContraste(Mat ImIn);

int main()
{
    Mat img, img_o1, img_o2, img_o3;
    int pos = 0;
    string* strMatricula = new string[8];
    Leer_Archivo();
    vector<Mat> HSV;
    VideoCapture cap;
    cap.open(1);
    cap.set(CV_CAP_PROP_FRAME_WIDTH,320);
    cap.set(CV_CAP_PROP_FRAME_HEIGHT,240);
    if(!cap.isOpened())
        return -1;
    namedWindow("webcam");
    namedWindow("out");
    namedWindow("DNI");

    while(1)
    {
        cap.read(img);
        cvtColor(img, img_o1,CV_BGR2HSV);
        Scalar hsv_l(70, 0, 100);
        Scalar hsv_h(120, 200, 255);
        Mat bw;
```

```

inRange(img_o1, hsv_l, hsv_h, bw);

imshow("webcam",bw);

CvPoint pt1[5000], pt2[5000];
// ENCONTRAMOS CONTORNOS DE POSIBLES DNI
vector<vector<Point> > contours;
vector<Vec4i> hierarchy;
bw.copyTo(img_o3);
findContours(img_o3, contours,
hierarchy,CV_RETR_EXTERNAL,CV_CHAIN_APPROX_SIMPLE);
CvRect rectP[5000];

for( int i = 0; i< contours.size(); i++ ) // FILTRAMOS POR AREA
{
    double ar=contourArea(contours[i],false);
    Scalar color(0,255,0);
    rectP[i] = boundingRect(contours[i]);
    pt1[i].x = rectP[i].x-10;
    if (pt1[i].x < 0) ar = 50;
    pt1[i].y = rectP[i].y-7;
    if (pt1[i].y < 0) ar = 50;
    pt2[i].x = (rectP[i].x + rectP[i].width)+10;
    if (rectP[i].x + rectP[i].width + 10 > bw.cols) ar = 50;
    pt2[i].y = (rectP[i].y + rectP[i].height)+7;
    if (rectP[i].y + rectP[i].height + 7 > bw.rows ) ar = 50;

    double Asp = rectP[i].width / rectP[i].height;
    // FILTRAMOS DNI POR AREA, RELACION ANCHO
    if(ar > 10000 && Asp > 0.9 && Asp < 1.5)//revisar
    {
        rectangle(img, pt1[i], pt2[i], Scalar(0,255,255), 2, 8, 0);
        // Extraer ROI DNI
    }
}

```



```

Mat roiDNI = img( Rect(pt1[i],pt2[i]));

roiDNI = mejorarContraste(roiDNI);

morphologyEx(roiDNI, roiDNI, MORPH_BLACKHAT,
getStructuringElement(MORPH_RECT,Size(5,1)));
threshold( roiDNI, roiDNI, 100.0,
255,CV_THRESH_BINARY);
imshow("DNI",roiDNI);
// Encontrar Número

int* histH = new int[roiDNI.cols];
int* histV = new int[roiDNI.rows];
for(int k = 0; k < roiDNI.cols; k++)
    histH[k] = 0;
for(int k = 0; k < roiDNI.rows; k++)
    histV[k] = 0;

for( int k = 0; k < roiDNI.rows; k++ )
{
    for(int ia = 50; ia < roiDNI.cols-50; ia++)
        histV[k] = histV[k] +
roiDNI.at<uchar>(k,ia);

    histV[k] /= 250;
}

// -----
Mat histPlotV( roiDNI.rows, 250, CV_8UC3 );
histPlotV = Scalar::all(0);
for(int k = 1; k < roiDNI.rows; k++)
    line(histPlotV,Point(histV[k-
1],k),Point(histV[k],k),Scalar(0,255,0));

```

```

int cont = 0, bInf = 0, bSup = 0, bIzq = 0, bDer =
roiDNI.cols;

pos = 0;
for(int k = roiDNI.rows-5; k > (roiDNI.rows/2); k--)
{
    if((histV[k-5] - histV[k]) > 30 && histV[k] < 20
&& histV[k-1] > 20)
    {
        line(histPlotV,Point(0,
k),Point(histPlotV.cols,k),Scalar(255,255,255));
        k = k - 3;
        cont++;
        bSup = k + 5;
    }

    if (cont == 3)
    {
        if((histV[k+5] - histV[k]) > 30 && histV[k]
< 20 && histV[k+1] > 20)
        {
            line(histPlotV,Point(0,
k),Point(histPlotV.cols,k),Scalar(255,0,255));

            cont = 0;
            bInf = k - 5;

            int imgX1 = pt1[i].x, imgY1 =
pt1[i].y, imgX2 = pt2[i].x, imgY2 = pt2[i].y;

            // Extracción a partir de la imagen
original

            //Mat roiNUM = img(
Rect(imgX1+0, imgY1+bInf, roiDNI.cols, bSup-bInf));

```

DNI

bInf, roiDNI.cols, bSup-bInf));

DE CARACTERES

contoursNum;

contoursNum, hierarchyNum, CV_RETR_EXTERNAL,
CV_CHAIN_APPROX_SIMPLE);

COORDENADAS X

id_Letra[100];

contoursNum.size(); id++)

boundingRect(contoursNum[id]);

rectPNum[id].x; // almacenamos coordenada X

// Extracción a partir de la región del

Mat roiNUM = roiDNI(Rect(0,

pos = 1;

namedWindow("roiNUM",1);

imshow("roiNUM",roiNUM);

// ENCONTRAMOS CONTORNOS

vector<vector<Point> >

vector<Vec4i> hierarchyNum;

Mat imgNum;

roiNUM.copyTo(imgNum);

findContours(imgNum,

CvRect rectPNum[1000];

// ALMACENAMOS

int coord_x_Letra[100],

for(int id = 0; id <

{

rectPNum[id] =

coord_x_Letra[id] =

id_Letra[id] = id;

```

    }

    int aux_p;

    for(int ii = 0; ii <
        contoursNum.size()-1; ii++)

        {
            for(int jj = ii+1; jj <
                contoursNum.size(); jj++)

                {
                    if(coord_x_Letra[ii] >
                        coord_x_Letra[jj])

                        {
                            aux_p =
                                id_Letra[ii];

                            id_Letra[ii] =
                                id_Letra[jj];

                            id_Letra[jj] =
                                aux_p;

                            coord_x_Letra[ii];

                            coord_x_Letra[ii] = coord_x_Letra[jj];

                            coord_x_Letra[jj] = aux_p;

                        }
                }
        }

    int c = 0, contNum = 0;

    for( int id = 0; id <
        contoursNum.size(); id++ )

```

```

{
    double ard =
contourArea(contoursNum[id_Letra[id]], false);
    if ( ard > 20 && ard < 300 )
    {

        rectPNum[id_Letra[id]] = boundingRect(contoursNum[id_Letra[id]]);

        Mat c1P(25, 15,
CV_8UC1 );

        Size size(15,25);
        Mat c1 =
roiNUM(rectPNum[id_Letra[id]]);

        resize(c1,
c1P,size,0,0,CV_INTER_CUBIC);

        threshold( c1P, c1P,
100.0, 255,CV_THRESH_BINARY);

        strMatricula[0] =
Reconocer(c1P, 0);

        rectPNum[id_Letra[id]].x = rectPNum[id_Letra[id]].x + imgX1;

        rectPNum[id_Letra[id]].y = rectPNum[id_Letra[id]].y + imgY1 + bInf;
        rectangle(img,
rectPNum[id_Letra[id]], Scalar(0,255,0) );

        //putText(img,
strMatricula[0], Point(rectPNum[id].x, rectPNum[id].y-5),
FONT_ITALIC,0.7,Scalar(0,0,255),2);

        if(c == 3)
        {

```

```

        strMatricula[0] = Reconocer(c1P, 1);

                                                                    contNum++;
                                                                    putText(img,
strMatricula[0], Point(pt1[i].x+(contNum*18), pt1[i].y-10),
FONT_ITALIC,0.7,Scalar(0,0,255),2);

                                                                    //putText(img,
"OK", Point(pt1[i].x, pt1[i].y-5), FONT_ITALIC,0.7,Scalar(0,0,255),2);
                                                                    if(contNum
== 8)

                                                                    {
                                                                    c = 0;
                                                                    break;
                                                                    }
                                                                    }

                                                                    if( strMatricula[0] ==
"R" && c == 2) c = 1;

                                                                    //putText(img,
"OK", Point(pt1[i].x, pt1[i].y-5), FONT_ITALIC,0.7,Scalar(0,0,255),2);

                                                                    //                                                                    //putText(img,
to_string(rectPNum[id].x), Point(rectPNum[id].x, rectPNum[id].y+20),
FONT_ITALIC,0.3,Scalar(0,0,255),1);

                                                                    //                                                                    putText(img,
to_string(id), Point(rectPNum[id_Letra[id]].x, rectPNum[id_Letra[id]].y+20),
FONT_ITALIC,0.3,Scalar(0,0,255),1);

                                                                    }
                                                                    }

```

```

                                break;
                            }
                        }
                    }
                namedWindow("HistV",1);
                imshow("HistV",histPlotV);

            }

        }

    imshow("out",img);

    if(waitKey(10) == 27) break;        // ESC
}
return 0;
}

```

ANEXO 2: RESULTADOS DE LECTURA DE IMÁGENES IMPRESAS DE DNI DE INTERNET



